

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problems Mailbox.**

**THIS PAGE BLANK (USPTO)**

# SEARCH REQUEST FORM

Requestor's Name: \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Date: \_\_\_\_\_ Phone: \_\_\_\_\_ Art Unit: \_\_\_\_\_

## Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

## STAFF USE ONLY

Date completed: 04-22-02  
Searcher: Beverly C 4997  
Terminal time: 20  
Elapsed time: \_\_\_\_\_  
CPU time: \_\_\_\_\_  
Total time: 25  
Number of Searches: \_\_\_\_\_  
Number of Databases: 1

### Search Site

\_\_\_\_ STIC  
\_\_\_\_ CM-1  
\_\_\_\_ Pre-S

### Type of Search

\_\_\_\_ N.A. Sequence  
\_\_\_\_ A.A. Sequence  
\_\_\_\_ Structure  
\_\_\_\_ Bibliographic

### Vendors

\_\_\_\_ IG  
\_\_\_\_ STN  
\_\_\_\_ Dialog  
\_\_\_\_ APS  
\_\_\_\_ Geninfo  
\_\_\_\_ SDC  
\_\_\_\_ DARC/Questel  
\_\_\_\_ Other CGN

**THIS PAGE BLANK (USPTO)**



STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,261  
FILING DATE:  
CLASSIFICATION: 435  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 137 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-256-261-14

Query Match  
Best Local Similarity 97.1%; Score 692; DB 1; Length 137;  
Matches 132; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSFDKGRKKGGDASYEPTGPTLVNVTGVDSKGNELLSPHYVEPIKPGTTLTKRKI 60  
DB 2 SSSFDKGRKKGGDASYEPTGPTLVNVTGVDSKGNELLSPHYVEPIKPGTTLTKRKI 61  
QY 61 EYVEWALDAPATYKKEFRVVELDPSAKIEVYTYDKNKKKEETKSPPIKEKGVVPLDSEHI 120  
DB 62 EYVEWALDAPATYKKEFRVVELDPSAKIEVYTYDKNKKKEETKSPPIKEKGVVPLDSEHI 121  
QY 121 KNPGFNLTTRKVIIEKK 136  
DB 122 KNPGFNLTTRKVIIEKK 137

RESULT 13  
US-08-852-299-14  
Sequence 14, Application US/08852299  
Patent No. 6010897

GENERAL INFORMATION:  
APPLICANT: Behnke, Detlef  
APPLICANT: Schlott, Bernhard  
APPLICANT: Albrecht, Sybille  
APPLICANT: G hrs, Karl-Heinz  
APPLICANT: Hartmann, Manfred  
TITLE OF INVENTION: Expression of signal-peptide-free  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESS: Fish & Neave  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/852,299  
FILING DATE: 17-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/256,261  
FILING DATE:  
INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:  
LENGTH: 137 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-852-299-14

Query Match  
Best Local Similarity 96.8%; Score 692; DB 3; Length 137;  
Matches 132; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSFDKGRKKGGDASYEPTGPTLVNVTGVDSKGNELLSPHYVEPIKPGTTLTKRKI 60  
DB 2 SSSFDKGRKKGGDASYEPTGPTLVNVTGVDSKGNELLSPHYVEPIKPGTTLTKRKI 61  
QY 61 EYVEWALDAPATYKKEFRVVELDPSAKIEVYTYDKNKKKEETKSPPIKEKGVVPLDSEHI 120  
DB 62 EYVEWALDAPATYKKEFRVVELDPSAKIEVYTYDKNKKKEETKSPPIKEKGVVPLDSEHI 121  
QY 121 KNPGFNLTTRKVIIEKK 136  
DB 122 KNPGFNLTTRKVIIEKK 137

RESULT 14  
US-08-256-261-12  
Sequence 12, Application US/08256261  
Patent No. 5801037

GENERAL INFORMATION:  
APPLICANT: Behnke, Detlef  
APPLICANT: Schlott, Bernhard  
APPLICANT: Albrecht, Sybille  
APPLICANT: G hrs, Karl-Heinz  
APPLICANT: Hartmann, Manfred  
TITLE OF INVENTION: Expression of signal-peptide-free  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESS: Fish & Neave  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,261  
FILING DATE:  
CLASSIFICATION: 435  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 137 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-256-261-12

Query Match  
Best Local Similarity 96.4%; Score 689; DB 1; Length 137;  
Matches 132; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSSFDKGRKKGGDASYEPTGPTLVNVTGVDSKGNELLSPHYVEPIKPGTTLTKRKI 60  
DB 2 SSSFDKGRKKGGDASYEPTGPTLVNVTGVDSKGNELLSPHYVEPIKPGTTLTKRKI 61  
QY 61 EYVEWALDAPATYKKEFRVVELDPSAKIEVYTYDKNKKKEETKSPPIKEKGVVPLDSEHI 120

Db 62 EYVEWALDATAKEFRVVDPSAKIEVTYDKNKKKEETKSPITEKGFVVPDLSEHI 121  
QY 121 KNPGFNLTIKVVEKK 136  
|||||  
Db 122 KNPGFNLTIKVVEKK 137

RESULT 15  
US-08-852-299-12  
; Sequence 12, Application US/08852299  
; Patent No. 6010897  
; GENERAL INFORMATION:  
; APPLICANT: Behnke, Detlef  
; APPLICANT: Schlott, Bernhard  
; APPLICANT: Albrecht, Sybille  
; APPLICANT: G hrs, Karl-Heinz  
; APPLICANT: Hartmann, Manfred  
; TITLE OF INVENTION: Expression of signal-peptide-free  
; TITLE OF INVENTION: staphylokinases  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Neave  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/852,299  
; FILING DATE: 17-MAY-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/256,261  
; FILING DATE:  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 137 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-852-299-12

Query Match 96.4%; Score 689; DB 3; Length 137;  
Best Local Similarity 97.1%; Pred. No. 1.9e-75;  
Matches 132; Conservative 0; Mismatches 4; Indels 0; Gaps  
QY 1 SSSFDKGYKGGDDASYFEPTGPIVMVNTGVDKGNELLSPHYVEFPKPGTTLTKKI 60  
|||||  
Db 2 SSSFDKGYKGGDDASYFEPTGPIVMVNTGVDKGNELLSPHYVEFPKPGTTLTKKI 61  
|||||  
QY 61 EYVEWALDATAKEFRVVDPSAKIEVTYDKNKKKEETKSPITEKGFVVPDLSEHI 120  
|||||  
Db 62 EYVEWALDATAKEFRVVDPSAKIEVTYDKNKKKEETKSPITEKGFVVPDLSEHI 121  
|||||  
QY 121 KNPGFNLTIKVVEKK 136  
|||||  
Db 122 KNPGFNLTIKVVEKK 137

Search completed: April 22, 2002, 10:41:49  
Job time: 90 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 10:41:55 ; Search time 23.85 Seconds

(without alignments)  
422.389 Million cell updates/sec

Title: us-09-601-490-1

Perfect score: 136  
Sequence: 1 SSSFDKGRKTKGGDASYPF.....SEHIKPNGENLTKVIEKK 136

Scoring table:  
OLIGO  
Gapox 60.0 , Gapext 60.0

Searched: 522463 seqs, 74073290 residues

Word size : 0

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

\_A\_Geneseq\_1101.\*  
1: /SID58/gcgdata/geneseq/geneseq/AA1980.DAT:\*  
2: /SID58/gcgdata/geneseq/geneseq/AA1981.DAT:\*  
3: /SID58/gcgdata/geneseq/geneseq/AA1982.DAT:\*  
4: /SID58/gcgdata/geneseq/geneseq/AA1983.DAT:\*  
5: /SID58/gcgdata/geneseq/geneseq/AA1984.DAT:\*  
6: /SID58/gcgdata/geneseq/geneseq/AA1985.DAT:\*  
7: /SID58/gcgdata/geneseq/geneseq/AA1986.DAT:\*  
8: /SID58/gcgdata/geneseq/geneseq/AA1987.DAT:\*  
9: /SID58/gcgdata/geneseq/geneseq/AA1988.DAT:\*  
10: /SID58/gcgdata/geneseq/geneseq/AA1989.DAT:\*  
11: /SID58/gcgdata/geneseq/geneseq/AA1990.DAT:\*  
12: /SID58/gcgdata/geneseq/geneseq/AA1991.DAT:\*  
13: /SID58/gcgdata/geneseq/geneseq/AA1992.DAT:\*  
14: /SID58/gcgdata/geneseq/geneseq/AA1993.DAT:\*  
15: /SID58/gcgdata/geneseq/geneseq/AA1994.DAT:\*  
16: /SID58/gcgdata/geneseq/geneseq/AA1995.DAT:\*  
17: /SID58/gcgdata/geneseq/geneseq/AA1996.DAT:\*  
18: /SID58/gcgdata/geneseq/geneseq/AA1997.DAT:\*  
19: /SID58/gcgdata/geneseq/geneseq/AA1998.DAT:\*  
20: /SID58/gcgdata/geneseq/geneseq/AA2000.DAT:\*  
21: /SID58/gcgdata/geneseq/geneseq/AA2001.DAT:\*  
22: /SID58/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	136	100.0	136	17	AAW03071
2	136	100.0	136	17	AAW03086
3	136	100.0	136	22	AAB93421
4	136	100.0	136	22	AAB61909
5	136	100.0	137	14	AAR39150
6	133	97.8	136	17	AAW03102
7	133	97.8	136	17	AAW03085
8	131	96.3	136	20	AAV15024
9	131	96.3	136	20	AAV15025
10	131	96.3	136	20	AAV15026
11	131	96.3	136	20	AAV15027

12	131	96.3	136	20	AAV15016	Staphylokinase var
13	130	95.6	136	17	AAW03087	Staphylokinase der
14	129	94.9	136	17	AAW03101	Staphylokinase der
15	129	94.9	136	20	AAV15023	Staphylokinase var
16	129	94.9	136	20	AAV15028	Staphylokinase var
17	129	94.9	136	20	AAV15015	Staphylokinase var
18	126	92.6	136	17	AAW03076	Staphylokinase der
19	126	92.6	136	17	AAW03088	Staphylokinase der
20	118	86.8	136	17	AAW03100	Staphylokinase der
21	118	86.8	136	17	AAW03084	Staphylokinase der
22	114	83.8	136	17	AAW03083	Staphylokinase der
23	114	83.8	136	17	AAW03099	Staphylokinase der
24	107	78.7	136	17	AAW03082	Staphylokinase der
25	107	78.7	136	17	AAW03098	Staphylokinase der
26	102	75.0	126	13	AAR25468	Thrombolytic Pepti
27	102	75.0	136	19	AAW44689	Staphylokinase mut
28	102	75.0	136	19	AAW44696	Staphylokinase fro
29	102	75.0	136	19	AAW44692	Staphylokinase mut
30	102	75.0	136	22	AAB61908	S. aureus staphylo
31	102	75.0	137	14	AAR39149	Staphylokinase SAK
32	102	75.0	163	12	AAR12137	S. aureus Staphylo
33	102	75.0	163	13	AAR28844	Staphylokinase (SAK)
34	101	74.3	136	17	AAW03077	Staphylokinase der
35	101	74.3	136	17	AAW03089	Staphylokinase der
36	101	74.3	136	20	AAV15021	Staphylokinase var
37	101	74.3	136	20	AAV15010	Staphylokinase var
38	100	73.5	136	20	AAV15018	Staphylokinase var
39	100	73.5	136	21	AAV98594	Staphylococcus pla
40	100	73.5	136	21	AAV50871	Staphylococcus aur
41	98	72.1	136	17	AAW03073	Staphylokinase der
42	98	72.1	136	17	AAW03090	Staphylokinase der
43	98	72.1	136	17	AAW03028	Staphylokinase der
44	95	69.9	136	17	AAW03097	Staphylokinase der
45	95	69.9	136	17	AAW03081	Staphylokinase der
46	94	69.1	136	20	AAV15047	Staphylokinase var
47	93	68.4	123	14	AAR39152	Staphylokinase SAK
48	93	68.4	127	14	AAR39151	Staphylokinase SAK
49	93	68.4	136	20	AAV15025	Staphylokinase var
50	93	68.4	136	20	AAV15009	Staphylokinase var
51	93	68.4	136	20	AAV15019	Staphylokinase var
52	93	68.4	136	20	AAV01232	S. aureus staphylo
53	93	68.4	137	14	AAR39154	Staphylokinase SAK
54	93	68.4	137	14	AAR39148	Staphylokinase SAK
55	93	68.4	137	14	AAR38153	Staphylokinase SAK
56	93	68.4	163	14	AAR39155	Staphylokinase SAK
57	93	68.4	163	16	AAR75207	Recombinant staphy
58	92	67.6	136	17	AAW03080	Staphylokinase der
59	92	67.6	136	17	AAW03096	Staphylokinase der
60	92	67.6	136	21	AAB01297	Wild type staphylo
61	91	66.9	136	20	AAV15014	Staphylokinase var
62	90	66.2	136	19	AAW44693	Staphylokinase mut
63	89	65.4	136	20	AAV15013	Staphylokinase var
64	85	62.5	136	17	AAW03079	Staphylokinase der
65	85	62.5	136	17	AAW03095	Staphylokinase der
66	81	59.6	136	20	AAV15007	Staphylokinase var
67	81	59.6	136	20	AAV15011	Staphylokinase var
68	81	59.6	136	20	AAV15012	Staphylokinase var
69	79	58.1	136	17	AAW03074	Staphylokinase der
70	79	58.1	136	17	AAW03094	Staphylokinase der
71	79	58.1	136	17	AAW03029	Staphylokinase der
72	79	58.1	136	20	AAV15052	Staphylokinase var
73	79	58.1	136	20	AAV15005	Staphylokinase var
74	79	58.1	136	20	AAV15006	Staphylokinase var
75	78	57.4	136	19	AAW44690	Staphylokinase mut
76	77	56.6	136	17	AAW03091	Staphylokinase der
77	74	54.4	136	19	AAW44695	Staphylokinase mut
78	74	54.4	136	19	AAV15008	Staphylokinase var
79	73	53.7	136	17	AAW03072	Staphylokinase der
80	73	53.7	136	17	AAW03093	Staphylokinase der
81	73	53.7	136	17	AAW03104	Staphylokinase der
82	73	53.7	136	17	AAW03027	Staphylokinase der
83	73	53.7	136	17	AAW03031	Staphylokinase der
84	73	53.7	136	20	AAV15066	Staphylokinase var

85	73	53.7	136	20	AAV15031	Staphylokinase var	158	15	11.0	15	21	AB37320	Peptide linker #5.
86	73	53.7	136	20	AAV15042	Staphylokinase var	159	15	11.0	15	22	AB61910	S. aureus staphylo
87	73	53.7	136	20	AAV15037	Staphylokinase var	160	15	11.0	15	22	AB99435	Staphylokinase min
88	73	53.7	136	20	AAV15022	Staphylokinase var	161	15	11.0	15	22	AB99435	Staphylokinase min
89	73	53.7	136	20	AAV15022	Staphylokinase var	162	14	10.3	17	7	AA61715	Fragment of sak ge
90	73	53.7	136	20	AAV15039	Staphylokinase var	163	13	9.6	17	22	AB99423	Staphylokinase min
91	73	53.7	136	20	AAV15041	Staphylokinase var	164	13	9.6	17	22	AB99434	Staphylokinase min
92	73	53.7	136	20	AAV15030	Staphylokinase var	165	12	8.8	12	15	AB67028	Fibrinolytic pepti
93	73	53.7	136	20	AAV15053	Staphylokinase var	166	12	8.8	12	15	AB67028	Fibrinolytic pepti
94	73	53.7	136	20	AAV15002	Staphylokinase var	167	12	8.8	12	15	AB67028	Fibrinolytic pepti
95	71	52.2	136	17	AAW03078	Staphylokinase var	168	12	8.8	12	15	AB67028	Fibrinolytic pepti
96	71	52.2	136	17	AAW03078	Staphylokinase var	169	12	8.8	12	15	AB67028	Fibrinolytic pepti
97	71	52.2	136	17	AAW03092	Staphylokinase der	170	12	8.8	12	15	AB67028	Fibrinolytic pepti
98	67	49.3	136	19	AAW44694	Staphylokinase var	171	12	8.8	12	15	AB67028	Fibrinolytic pepti
99	64	47.1	136	20	AAV15063	Staphylokinase mut	172	11	8.1	11	22	AB99433	Staphylokinase min
100	64	47.1	136	20	AAV15064	Staphylokinase var	173	11	8.1	11	22	AB99411	Staphylokinase min
101	64	47.1	136	20	AAV15067	SakSTAR (E65D, K74	174	11	8.1	11	22	AB99412	Staphylokinase T-C
102	64	47.1	136	20	AAV15068	SakSTAR (E65D, K74	175	11	8.1	11	22	AB99414	Staphylokinase T-C
103	64	47.1	136	20	AAV15069	Staphylokinase var	176	11	8.1	11	22	AB99415	Staphylokinase T-C
104	64	47.1	136	20	AAV15056	Staphylokinase var	177	11	8.1	11	22	AB99416	Staphylokinase T-C
105	64	47.1	136	20	AAV15043	Staphylokinase var	178	11	8.1	11	22	AB99417	Staphylokinase T-C
106	64	47.1	136	20	AAV15044	Staphylokinase var	179	11	8.1	11	22	AB99418	Staphylokinase T-C
107	64	47.1	136	20	AAV15057	Staphylokinase var	180	11	8.1	11	22	AB99419	Staphylokinase T-C
108	64	47.1	136	20	AAV15045	Staphylokinase var	181	11	8.1	11	22	AB99420	Staphylokinase T-C
109	64	47.1	136	20	AAV15046	Staphylokinase var	182	11	8.1	11	22	AB99437	Staphylokinase T-C
110	64	47.1	136	20	AAV15048	Staphylokinase var	183	11	8.1	11	22	AB99438	Staphylokinase T-C
111	64	47.1	136	20	AAV15059	Staphylokinase var	184	11	8.1	11	22	AB99439	Staphylokinase T-C
112	64	47.1	136	20	AAV15049	Staphylokinase var	185	11	8.1	11	22	AB99440	Staphylokinase T-C
113	64	47.1	136	20	AAV15050	Staphylokinase var	186	11	8.1	11	22	AB99441	Staphylokinase T-C
114	64	47.1	136	20	AAV15050	Staphylokinase var	187	11	8.1	11	22	AB99442	Staphylokinase T-C
115	64	47.1	136	20	AAV15061	Staphylokinase var	188	11	8.1	11	22	AB99443	Staphylokinase T-C
116	64	47.1	136	20	AAV15058	Staphylokinase var	189	11	8.1	11	22	AB99444	Staphylokinase T-C
117	64	47.1	136	20	AAV15040	Staphylokinase var	190	11	8.1	11	22	AB99445	Staphylokinase T-C
118	64	47.1	136	20	AAV15035	Staphylokinase var	191	11	8.1	11	22	AB99446	Staphylokinase T-C
119	64	47.1	136	20	AAV15036	Staphylokinase var	192	11	8.1	11	22	AB99447	Staphylokinase T-C
120	64	47.1	136	20	AAV15051	SakSTAR (E65D, K74	193	11	8.1	11	22	AB99448	Staphylokinase T-C
121	64	47.1	136	8	AAV70381	Sequence encoded b	194	11	8.1	11	22	AB99449	Staphylokinase T-C
122	62	45.6	136	20	AAV15020	Staphylokinase var	195	11	8.1	11	22	AB99450	Staphylokinase T-C
123	61	44.9	136	20	AAV15078	Staphylokinase var	196	11	7.4	15	21	AB99432	Staphylokinase min
124	61	44.9	136	20	AAV15079	Staphylokinase var	197	10	7.4	15	21	AB99432	Staphylokinase min
125	61	44.9	136	20	AAV15080	Staphylokinase var	198	10	7.4	15	21	AB99432	Staphylokinase min
126	61	44.9	136	20	AAV15076	Staphylokinase var	199	9	6.6	12	15	AB99427	Staphylokinase min
127	61	44.9	136	20	AAV15077	Staphylokinase var	200	9	6.6	12	15	AB99428	Staphylokinase min
128	61	44.9	136	20	AAV15077	Staphylokinase var	201	9	6.6	12	15	AB99428	Staphylokinase min
129	59	43.4	136	17	AAW03075	Staphylokinase der	202	8	5.9	17	22	AB99430	Staphylokinase min
130	59	43.4	136	17	AAW03103	Staphylokinase der	203	8	5.9	17	22	AB99430	Staphylokinase min
131	59	43.4	136	17	AAW03030	Staphylokinase der	204	7	5.1	7	22	AB99442	Staphylokinase min
132	56	41.2	136	20	AAV15062	SakSTAR (K57A, E58	205	7	5.1	17	22	AB99406	Staphylokinase min
133	55	40.4	136	20	AAV15058	Staphylokinase var	206	7	5.1	105	22	AAV16576	Staphylokinase min
134	55	40.4	136	20	AAV15034	Staphylokinase var	207	7	5.1	105	22	AAV16576	Staphylokinase min
135	55	40.4	136	20	AAV15032	Staphylokinase var	208	7	5.1	105	22	AAV16576	Staphylokinase min
136	55	40.4	136	20	AAV15033	Staphylokinase var	209	7	5.1	126	21	AAV1785	Staphylokinase min
137	51	37.5	136	19	AAW44691	Staphylokinase var	210	7	5.1	126	21	AAV1785	Staphylokinase min
138	48	35.3	109	7	AAV60796	Fragment of sak ge	211	7	5.1	177	20	AAV34220	Staphylokinase min
139	47	34.6	136	20	AAV15065	SakSTAR (K35A, E65	212	7	5.1	229	21	AAV37667	Staphylokinase min
140	47	34.6	136	20	AAV15070	Staphylokinase var	213	7	5.1	230	15	AAV34219	Staphylokinase min
141	47	34.6	136	20	AAV15071	SakSTAR (K35A, E65	214	7	5.1	230	15	AAV34219	Staphylokinase min
142	47	34.6	136	20	AAV15055	Staphylokinase var	215	7	5.1	295	21	AAV34218	Staphylokinase min
143	34	25.0	136	20	AAV15072	Staphylokinase var	216	7	5.1	372	22	AAV34218	Staphylokinase min
144	34	25.0	136	20	AAV15001	Staphylokinase gen	217	7	5.1	372	22	AAV34218	Staphylokinase min
145	33	24.3	158	12	AAV12138	Staphylokinase-50m	218	7	5.1	506	21	AAV1649	Staphylokinase min
146	31	22.8	136	20	AAV15073	Staphylokinase-50m	219	7	5.1	716	22	AAV1649	Staphylokinase min
147	31	22.8	136	20	AAV15074	Staphylokinase var	220	7	5.1	1014	9	AAV3126	Staphylokinase min
148	31	22.8	136	20	AAV15075	Staphylokinase var	221	7	5.1	1014	9	AAV3126	Staphylokinase min
149	28	20.6	135	21	AAV01298	Altered staphyloki	222	7	5.1	1027	11	AAV1773	Staphylokinase min
150	17	12.5	17	22	AAV99400	Staphylokinase var	223	7	5.1	1035	11	AAV1773	Staphylokinase min
151	17	12.5	17	22	AAV99401	Staphylokinase var	224	7	5.1	1036	13	AAV20559	Staphylokinase min
152	17	12.5	17	22	AAV99402	Staphylokinase var	225	7	5.1	1036	13	AAV20559	Staphylokinase min
153	17	12.5	17	22	AAV99403	Staphylokinase var	226	7	5.1	1036	13	AAV20559	Staphylokinase min
154	17	12.5	17	22	AAV99404	Staphylokinase var	227	7	5.1	1036	13	AAV20559	Staphylokinase min
155	17	12.5	17	22	AAV99405	Staphylokinase var	228	7	5.1	1409	21	AAV3125	Staphylokinase min
156	16	11.8	17	22	AAV99422	Staphylokinase var	229	6	4.4	1420	21	AAV3125	Staphylokinase min
157	16	11.8	17	22	AAV99436	Staphylokinase min	230	6	4.4	1420	21	AAV3125	Staphylokinase min

231	6	4.4	9	21	AAV86687	Telomerase peptide	304	6	4.4	231	22	AAV86902	Neisseria meningit
232	6	4.4	9	21	AAV86713	Telomerase peptide	305	6	4.4	231	22	AAV86903	Neisseria meningit
233	6	4.4	9	21	AAV86819	Telomerase peptide	306	6	4.4	231	22	AAV86904	Neisseria meningit
234	6	4.4	10	21	AAV96394	HIV-A2 binding pep	307	6	4.4	231	22	AAV86905	Neisseria meningit
235	6	4.4	13	18	AAV66936	Peptide epitope 29	308	6	4.4	231	22	AAV86906	Neisseria meningit
236	6	4.4	14	18	AAV66936	Peptide epitope 29	309	6	4.4	231	22	AAV86907	Neisseria meningit
237	6	4.4	17	18	AAV66936	Peptide epitope 81	310	6	4.4	231	22	AAV86908	Neisseria meningit
238	6	4.4	17	18	AAV66936	Peptide epitope 81	311	6	4.4	231	22	AAV86909	Neisseria meningit
239	6	4.4	18	18	AAV66914	Staphylokinase Sak	312	6	4.4	231	22	AAV86910	Neisseria meningit
240	6	4.4	18	18	AAV66914	Peptide epitope 9	313	6	4.4	231	22	AAV86911	Neisseria meningit
241	6	4.4	19	18	AAV66915	Peptide epitope 12	314	6	4.4	231	22	AAV86912	Neisseria meningit
242	6	4.4	19	18	AAV66915	Peptide epitope 10	315	6	4.4	231	22	AAV86913	Neisseria meningit
243	6	4.4	19	18	AAV66915	Peptide epitope 13	316	6	4.4	231	22	AAV86914	Neisseria meningit
244	6	4.4	31	21	AAV58831	Native mouse IgG1	317	6	4.4	235	18	AAV26692	Toxoplasma gondii
245	6	4.4	31	21	AAV58831	Arabisdopsis thalia	318	6	4.4	235	21	AAV16943	Toxoplasma gondii
246	6	4.4	32	21	AAV588091	Arabisdopsis thalia	319	6	4.4	236	21	AAV57749	Toxoplasma gondii
247	6	4.4	36	21	AAV42295	CD47 extracellular	320	6	4.4	241	22	AAV86458	S. pneumoniae muta
248	6	4.4	41	18	AAV66905	CD47 extracellular	321	6	4.4	244	21	AAV86458	Arabisdopsis thalia
249	6	4.4	51	22	AAV64883	Human secreted pro	322	6	4.4	247	21	AAV86296	Arabisdopsis thalia
250	6	4.4	62	21	AAV64883	Arabisdopsis thalia	323	6	4.4	250	21	AAV86222	Arabisdopsis thalia
251	6	4.4	63	22	AAV18293	Peptide #4727 enco	324	6	4.4	259	22	AAV86464	Hydroxyindol-relat
252	6	4.4	77	21	AAV80884	Human secreted pro	325	6	4.4	259	22	AAV862248	Rat insulin-respon
253	6	4.4	92	21	AAV80884	Arabisdopsis thalia	326	6	4.4	263	21	AAV805865	Arabisdopsis thalia
254	6	4.4	99	22	AAV89310	Human secreted pro	327	6	4.4	266	21	AAV81386	Human ORF ORF1150
255	6	4.4	104	21	AAV83028	Pinus radiata tran	328	6	4.4	272	18	AAV26698	Toxoplasma gondii
256	6	4.4	106	19	AAV86028	S. pneumoniae derl	329	6	4.4	272	21	AAV54420	Secoisolaricresin
257	6	4.4	135	21	AAV827690	Arabisdopsis thalia	330	6	4.4	273	21	AAV54412	Secoisolaricresin
258	6	4.4	135	21	AAV827690	Arabisdopsis thalia	331	6	4.4	276	21	AAV54416	Secoisolaricresin
259	6	4.4	139	19	AAV66427	Human secreted pro	332	6	4.4	277	18	AAV55605	H. pylori ORF 14ce
260	6	4.4	141	20	AAV55818	Amino end of the	333	6	4.4	277	21	AAV54413	Secoisolaricresin
261	6	4.4	141	21	AAV827689	Arabisdopsis thalia	334	6	4.4	277	21	AAV54415	Secoisolaricresin
262	6	4.4	141	21	AAV827689	Arabisdopsis thalia	335	6	4.4	278	16	AAV72597	Ostertaria protease
263	6	4.4	142	18	AAV66904	CD47 extracellular	336	6	4.4	287	22	AAV65782	Cysteine protease
264	6	4.4	146	21	AAV66904	Human secreted pro	337	6	4.4	287	22	AAV66221	Arabisdopsis thalia
265	6	4.4	149	21	AAV66904	Helicobacter pylori	338	6	4.4	293	18	AAV55218	H. pylori ORF 02ce
266	6	4.4	150	21	AAV66904	Plasmid pCHO-shlAP	339	6	4.4	294	22	AAV83176	Cornebacterium th
267	6	4.4	153	19	AAV87516	H. pylori GHPD 128	340	6	4.4	307	22	AAV81714	S. epidermidis ope
268	6	4.4	155	22	AAV87516	Human colon cancer	341	6	4.4	319	19	AAV72059	HSV-2 strain SB5
269	6	4.4	158	21	AAV82493	Helicobacter pylori	342	6	4.4	320	14	AAV83279	43 kd endoflagellu
270	6	4.4	164	21	AAV827688	Arabisdopsis thalia	343	6	4.4	320	21	AAV82855	Arabisdopsis thalia
271	6	4.4	164	21	AAV827688	Arabisdopsis thalia	344	6	4.4	321	20	AAV35152	Amino acid sequenc
272	6	4.4	165	21	AAV827688	Arabisdopsis thalia	345	6	4.4	324	22	AAV80292	C glutamyl prote
273	6	4.4	165	21	AAV827688	Arabisdopsis thalia	346	6	4.4	330	20	AAV88339	Salmonella enteric
274	6	4.4	175	17	AAV87507	Caenorhabditis ele	347	6	4.4	332	20	AAV88068	Streptococcus pyog
275	6	4.4	175	21	AAV87507	Human cancer assoc	348	6	4.4	335	21	AAV95073	Candida albicans p
276	6	4.4	178	21	AAV71064	Human membrane tra	349	6	4.4	335	22	AAV85002	Shrimp white spot
277	6	4.4	178	22	AAV831898	Amino acid sequenc	350	6	4.4	338	19	AAV80521	Kurtia sp. DTB sy
278	6	4.4	179	21	AAV831898	Arabisdopsis thalia	351	6	4.4	338	22	AAV80521	C glutamyl prote
279	6	4.4	189	22	AAV831900	Amino acid sequenc	352	6	4.4	345	15	AAV81848	Influenza type B H
280	6	4.4	191	21	AAV831900	Arabisdopsis thalia	353	6	4.4	345	15	AAV81848	Influenza type B H
281	6	4.4	191	21	AAV831900	Amino acid sequenc	354	6	4.4	345	15	AAV81848	Influenza type B H
282	6	4.4	193	22	AAV831896	Amino acid sequenc	355	6	4.4	345	15	AAV81850	Influenza type B H
283	6	4.4	193	22	AAV831896	Amino acid sequenc	356	6	4.4	345	15	AAV81851	Influenza type B H
284	6	4.4	193	22	AAV831903	Amino acid sequenc	357	6	4.4	345	15	AAV81852	Influenza type B H
285	6	4.4	193	22	AAV831903	Amino acid sequenc	358	6	4.4	345	15	AAV81853	Influenza type B H
286	6	4.4	193	22	AAV831904	Amino acid sequenc	359	6	4.4	345	15	AAV81854	Influenza type B H
287	6	4.4	196	21	AAV831928	Arabisdopsis thalia	360	6	4.4	345	15	AAV81856	Influenza type B H
288	6	4.4	200	21	AAV831899	Amino acid sequenc	361	6	4.4	345	15	AAV81857	Influenza type B H
289	6	4.4	202	21	AAV831899	Arabisdopsis thalia	362	6	4.4	345	15	AAV81859	Influenza type B H
290	6	4.4	213	21	AAV831899	Arabisdopsis thalia	363	6	4.4	347	15	AAV81845	Influenza type B H
291	6	4.4	221	22	AAV830556	A wild type rat 11	364	6	4.4	347	15	AAV81846	Influenza type B H
292	6	4.4	221	22	AAV830557	A mutagenised rat	365	6	4.4	347	15	AAV81847	Influenza type B H
293	6	4.4	221	22	AAV830558	A mutagenised rat	366	6	4.4	347	15	AAV81848	Influenza type B H
294	6	4.4	221	22	AAV830559	A mutagenised rat	367	6	4.4	351	21	AAV818970	Amino acid sequenc
295	6	4.4	221	22	AAV830560	A mutagenised rat	368	6	4.4	351	22	AAV83685	Human protein sequ
296	6	4.4	221	22	AAV830561	A mutagenised rat	369	6	4.4	353	22	AAV83685	Human colon cancer
297	6	4.4	221	22	AAV830562	A mutagenised rat	370	6	4.4	360	22	AAV830362	C glutamyl prote
298	6	4.4	221	22	AAV830563	A mutagenised rat	371	6	4.4	364	22	AAV839861	Human polypeptide
299	6	4.4	221	22	AAV830564	A mutagenised rat	372	6	4.4	364	22	AAV839861	Human protein sequ
300	6	4.4	221	22	AAV830565	A mutagenised rat	373	6	4.4	364	22	AAV839861	Corynebacterium gl
301	6	4.4	228	21	AAV57750	Toxoplasma gondii	374	6	4.4	365	21	AAV839861	Zea mays protein f
302	6	4.4	229	13	AAV81613	UCH-L3 protein. H	375	6	4.4	365	22	AAV839861	Human protein sequ
303	6	4.4	230	20	AAV839861	UCH-L3 protein. H	376	6	4.4	367	22	AAV839861	Human colon cancer

377	6	4.4	370	22	AAW41647	Human polypeptide	450	6	4.4	489	20	AAV26065	Para-Nitrobenzyl E
378	6	4.4	372	22	AAW81006	HIV protease and r	451	6	4.4	496	20	AAV73823	Human prostate tum
379	6	4.4	372	22	AAW81023	HIV protease and r	452	6	4.4	499	22	AAW40963	Human polypeptide
380	6	4.4	373	21	AAW81049	HIV protease and r	453	6	4.4	499	22	AAW82251	Rat insulin-respon
381	6	4.4	374	21	AAW21092	Modified Treponema	454	6	4.4	503	20	AAW00189	Enterococcus faeca
382	6	4.4	374	22	AAW83001	S. epidermidis ope	455	6	4.4	503	21	AAW25432	Arabisdopsis thalia
383	6	4.4	376	18	AAW32863	Saccharopolyspora	456	6	4.4	506	17	AAW95594	Arabisdopsis fatty
384	6	4.4	376	20	AAV32054	Bovine pregnancy a	457	6	4.4	506	21	AAV77551	Arabisdopsis thalia
385	6	4.4	398	12	AAW15614	Human type II inte	458	6	4.4	508	21	AAW18187	Arabisdopsis thalia
386	6	4.4	398	14	AAW42061	Lymphoblastoid der	459	6	4.4	513	20	AAV05897	Plasmodium falcipa
387	6	4.4	398	15	AAW60617	Human type II IL-1	460	6	4.4	515	21	AAW46671	Vicia sativa omega
388	6	4.4	398	16	AAW85480	Human type II IL-1	461	6	4.4	527	21	AAW48371	Arabisdopsis thalia
389	6	4.4	398	19	AAW59267	Human type II IL-1	462	6	4.4	535	21	AAW25431	Arabisdopsis thalia
390	6	4.4	398	21	AAW37792	Human Interleukin-	463	6	4.4	535	21	AAW28064	Arabisdopsis thalia
391	6	4.4	398	22	AAW24185	Human EST encoded	464	6	4.4	537	21	AAW44998	Arabisdopsis thalia
392	6	4.4	410	22	AAW50081	Human IL-1R type I	465	6	4.4	538	19	AAW47001	Guaiaciflone-S-tran
393	6	4.4	410	22	AAW86456	S. pneumoniae fibb	466	6	4.4	538	21	AAW44248	Arabisdopsis thalia
394	6	4.4	415	14	AAW32922	AAWP-1. Homo sapi	467	6	4.4	547	22	AAW50670	C. elegans UNC-5 P
395	6	4.4	415	16	AAW65773	N-terminally trunc	468	6	4.4	549	21	AAW14497	Arabisdopsis thalia
396	6	4.4	415	18	AAW35692	Treponema pallidum	469	6	4.4	551	19	AAW59333	Modified T. pallid
397	6	4.4	415	19	AAW68498	Treponema pallidum	470	6	4.4	554	21	AAW36165	Arabisdopsis thalia
398	6	4.4	415	19	AAW59930	Modified T. pallid	471	6	4.4	559	20	AAW16102	Acetobacter xylinu
399	6	4.4	426	16	AAW66776	Cell adhesion prot	472	6	4.4	585	18	AAW01671	Influenza B/Panama
400	6	4.4	434	19	AAW59934	Wild type T. palli	473	6	4.4	585	20	AAW75443	Influenza virus B/
401	6	4.4	440	21	AAW28066	Arabisdopsis thalia	474	6	4.4	585	22	AAE04953	Influenza virus B/
402	6	4.4	441	21	AAW14499	Arabisdopsis thalia	475	6	4.4	586	18	AAW01675	Influenza B/Hardin
403	6	4.4	443	15	AAW58637	Treponema pallidum	476	6	4.4	586	20	AAW75447	Influenza virus B/
404	6	4.4	443	18	AAW35744	Treponema pallidum	477	6	4.4	586	22	AAE04957	Influenza virus B/
405	6	4.4	454	21	AAW25433	Arabisdopsis thalia	478	6	4.4	589	18	AAW01672	Influenza B/Necker
406	6	4.4	456	18	AAW33693	Nucleic acid-bound	479	6	4.4	589	20	AAW75444	Influenza virus B/
407	6	4.4	457	21	AAW28065	Arabisdopsis thalia	480	6	4.4	589	22	AAE04954	Influenza virus B/
408	6	4.4	461	19	AAW70964	Human Ikaros isofo	481	6	4.4	591	20	AAW97384	A catalytic telome
409	6	4.4	468	21	AAW51824	Human OSBH protein	482	6	4.4	592	18	AAW01674	Influenza A/Shanha
410	6	4.4	470	19	AAW47172	Glucosyl transfera	483	6	4.4	592	20	AAW75446	Influenza virus B/
411	6	4.4	470	20	AAW06274	Tomato dehydridico	484	6	4.4	592	22	AAE04956	Influenza virus B/
412	6	4.4	476	21	AAW45086	Tobacco salicylic	485	6	4.4	596	21	AAW46670	Arabisdopsis thalia
413	6	4.4	476	21	AAW54581	Amino acid sequenc	486	6	4.4	603	22	AAW94738	Human protein sequ
414	6	4.4	479	21	AAW21009	Human nucleic acid	487	6	4.4	610	21	AAW18247	Plasmodium falcipa
415	6	4.4	479	22	AAW39177	Human polypeptide	488	6	4.4	612	22	AAW88482	Human membrane or
416	6	4.4	479	22	AAW93610	Lipid associated p	489	6	4.4	622	22	AAW49771	Amyloid-beta prote
417	6	4.4	480	21	AAW97284	Arabisdopsis thalia	490	6	4.4	622	20	AAW25463	Human CRT-1 protei
418	6	4.4	480	21	AAW29088	B. subtilis para-n	491	6	4.4	629	19	AAW80721	S. pneumoniae DNA
419	6	4.4	489	14	AAW38748	Para-nitrobenzyl-e	492	6	4.4	631	21	AAW46669	Arabisdopsis thalia
420	6	4.4	489	17	AAW85755	Para-nitrobenzyl e	493	6	4.4	652	18	AAW18010	Plasmodium falcipa
421	6	4.4	489	18	AAW2135	Modified para-nitr	494	6	4.4	652	21	AAW36164	Arabisdopsis thalia
422	6	4.4	489	18	AAW22137	Modified para-nitr	495	6	4.4	666	21	AAW79506	Human carbohydrate
423	6	4.4	489	18	AAW22138	Modified para-nitr	496	6	4.4	674	19	AAW36522	Human adducin-like
424	6	4.4	489	18	AAW22140	Modified para-nitr	497	6	4.4	686	21	AAW81293	Plasmodium falcipa
425	6	4.4	489	18	AAW22139	Modified para-nitr	498	6	4.4	713	15	AAW60101	Canine zona pelluc
426	6	4.4	489	18	AAW22141	Modified para-nitr	499	6	4.4	715	15	AAW55198	Canine zona pelluc
427	6	4.4	489	18	AAW22142	Modified para-nitr	500	6	4.4	715	20	AAW42471	Canine zona pelluc
428	6	4.4	489	18	AAW22143	Modified para-nitr	501	6	4.4	715	20	AAW81808	Canine zPA protein
429	6	4.4	489	18	AAW22144	Modified para-nitr	502	6	4.4	715	21	AAW82206	Canine zona pelluc
430	6	4.4	489	18	AAW22145	Modified para-nitr	503	6	4.4	715	21	AAW52680	Canine oocyte zona
431	6	4.4	489	18	AAW22145	Modified para-nitr	504	6	4.4	715	21	AAW52171	Canine zona pelluc
432	6	4.4	489	20	AAW26072	Para-Nitrobenzyl E	505	6	4.4	715	21	AAW52975	Canine zona pelluc
433	6	4.4	489	20	AAW26073	Para-Nitrobenzyl E	506	6	4.4	728	13	AAW20100	Canine alpha- an
434	6	4.4	489	20	AAW26074	Para-Nitrobenzyl E	507	6	4.4	733	21	AAW52241	Arabisdopsis thalia
440	6	4.4	489	20	AAW26238	Modified Para-Nitr	513	6	4.4	803	20	AAW97880	N-terminal domain
441	6	4.4	489	20	AAW26075	Para-Nitrobenzyl E	514	6	4.4	811	21	AAW52239	Arabisdopsis thalia
436	6	4.4	489	20	AAW26192	Para-Nitrobenzyl E	509	6	4.4	779	21	AAW52240	Arabisdopsis thalia
437	6	4.4	489	20	AAW26193	Para-Nitrobenzyl E	510	6	4.4	781	21	AAW36163	Arabisdopsis thalia
438	6	4.4	489	20	AAW26194	Para-Nitrobenzyl E	511	6	4.4	797	17	AAW86866	Human protocadheri
439	6	4.4	489	20	AAW26195	Para-Nitrobenzyl E	512	6	4.4	803	15	AAW60882	P. abyss DNA poly
440	6	4.4	489	20	AAW26238	Modified Para-Nitr	513	6	4.4	803	20	AAW97880	Maize phosphatidyl
441	6	4.4	489	20	AAW26075	Para-Nitrobenzyl E	514	6	4.4	811	21	AAW52239	Arabisdopsis thalia
442	6	4.4	489	20	AAW26063	Para-Nitrobenzyl E	515	6	4.4	813	15	AAW62754	Self sequence. Sa
443	6	4.4	489	20	AAW26064	Para-Nitrobenzyl E	516	6	4.4	814	18	AAW23573	Salmonella enterit
444	6	4.4	489	20	AAW26066	Para-Nitrobenzyl E	517	6	4.4	834	18	AAW11325	Human Na+/H+ excha
445	6	4.4	489	20	AAW26067	Para-Nitrobenzyl E	518	6	4.4	911	19	AAW74723	Human secreted pro
446	6	4.4	489	20	AAW26068	Para-Nitrobenzyl E	519	6	4.4	916	22	AAW38942	Human polypeptide
447	6	4.4	489	20	AAW26069	Para-Nitrobenzyl E	520	6	4.4	923	22	AAW38951	Human polypeptide
448	6	4.4	489	20	AAW26070	Para-Nitrobenzyl E	521	6	4.4	966	19	AAW49077	Solanum tuberosum
449	6	4.4	489	20	AAW26071	Para-Nitrobenzyl E	522	6	4.4	972	22	AAE01978	Human ATPase-relat

523	6	4.4	1008	22	AAB82247	Rat insulin-respon	596	5	3.7	8	18	AAW36925	Peptide epitope 20
524	6	4.4	1024	21	AAAG2350	Arabidopsis thalia	597	5	3.7	9	18	AAW36926	Peptide epitope 21
525	6	4.4	1036	21	AAAG1888	Arabidopsis thalia	598	5	3.7	9	20	AAAY40779	S6 derivative #11,
526	6	4.4	1055	16	AAW70114	IL-RTPeirt-GBP 13	599	5	3.7	9	21	AAW45561	Human B99-1 HLA B*
527	6	4.4	1055	22	AAW39198	Human polypeptide	600	5	3.7	9	21	AAW30117	Scaffold protein S
528	6	4.4	1056	22	AAE01980	Human ATPase-relat	601	5	3.7	9	21	AAW30118	Scaffold protein S
529	6	4.4	1070	21	AAAG31887	Arabidopsis thalia	602	5	3.7	9	21	AAW86786	Telomerase peptide
530	6	4.4	1074	20	AAAY00206	Enterococcus faeca	603	5	3.7	9	22	AAW22310	HIV peptide SEQ ID
531	6	4.4	1074	20	AAAY00188	Enterococcus faeca	604	5	3.7	10	20	AAW84235	Antigenic peptide
532	6	4.4	1096	21	AAAG48592	Arabidopsis thalia	605	5	3.7	10	20	AAW84236	Antigenic peptide
533	6	4.4	1096	21	AAAG48833	Amino acid sequenc	606	5	3.7	10	22	AAAG95753	Saccharomyces cere
534	6	4.4	1105	21	AAAG42349	Arabidopsis thalia	607	5	3.7	10	22	AAAG86101	Saccharomyces cere
535	6	4.4	1120	20	AAAY00641	Telomerase protein	608	5	3.7	10	22	AAW848575	Human MO-Bc-203 de
536	6	4.4	1120	20	AAAY00650	Telomerase (ver. 2	609	5	3.7	11	16	AAW21464	Schistosoma elasta
537	6	4.4	1130	21	AAAG48591	Arabidopsis thalia	610	5	3.7	11	21	AAW19278	A ricin toxin A ch
538	6	4.4	1132	19	AAW71376	Human telomerase c	611	5	3.7	11	22	AAW86304	Murine derived ant
539	6	4.4	1132	19	AAW46957	Human telomerase r	612	5	3.7	11	22	AAW50076	csk tyrosine kinas
540	6	4.4	1132	19	AAW56113	Human telomerase r	613	5	3.7	12	15	AAW67139	Factor VIIa deri
541	6	4.4	1132	20	AAAY36621	A human telomerase	614	5	3.7	12	18	AAW33289	Peptide 8 used in
542	6	4.4	1132	20	AAAY28861	Human telomerase r	615	5	3.7	12	22	AAW05218	Zinc finger protei
543	6	4.4	1132	20	AAAY32090	Human telomerase r	616	5	3.7	12	22	AAW49264	Zinc finger helica
544	6	4.4	1132	20	AAAY28401	Human EST2 protein	617	5	3.7	13	18	AAW99278	Peptide C1 from WO
545	6	4.4	1132	20	AAAY26580	Human telomerase r	618	5	3.7	13	18	AAW33288	Peptide 7 used in
546	6	4.4	1132	20	AAAY00627	Human telomerase p	619	5	3.7	13	15	AAW53343	Chimeric human Ab
547	6	4.4	1132	20	AAAY00638	Truncated telomera	620	5	3.7	14	18	AAW92277	Peptide C0 from WO
548	6	4.4	1132	21	AAW90251	Human catalytic te	621	5	3.7	14	18	AAW33287	Peptide 6 used in
549	6	4.4	1132	21	AAAY96566	hEST2, a human tel	622	5	3.7	14	21	AAW01630	Chimeric immunogl
550	6	4.4	1132	22	AAAB99330	Human telomerase p	623	5	3.7	14	21	AAAY65739	Breast cancer susc
551	6	4.4	1132	22	AAAG64329	Human protein #2.	624	5	3.7	15	10	AAAP91741	Antigenic C-termin
552	6	4.4	1132	22	AAAG64859	Heart muscle cell	625	5	3.7	15	14	AAW33253	PKM61HR3 derived
553	6	4.4	1140	22	AAW50669	C. elegans UNC-5 p	626	5	3.7	15	14	AAW36528	Heavy chain variab
554	6	4.4	1144	22	AAW59293	Human polypeptide	627	5	3.7	16	15	AAW45190	CS134-149 ("S1").
555	6	4.4	1150	19	AAW47006	Glutathione-S-tran	628	5	3.7	16	18	AAW40945	Cryptic peptide 2.
556	6	4.4	1154	19	AAW61350	Human telomerase p	629	5	3.7	16	20	AAW25535	Human MHC Class II
557	6	4.4	1165	22	AAW39312	Human polypeptide	630	5	3.7	17	21	AAW828948	Peptide Fel 33.. F
558	6	4.4	1166	20	AAAY00647	Telomerase (ver. 2	631	5	3.7	17	21	AAW87689	Feline human TRFP
559	6	4.4	1177	22	AAW41079	Human polypeptide	632	5	3.7	17	21	AAAY90119	Cat TRFP derived p
560	6	4.4	1177	22	AAW41098	Human polypeptide	633	5	3.7	17	21	AAW51486	Human TRFP derived
561	6	4.4	1189	19	AAW47008	Glutathione-S-tran	634	5	3.7	17	22	AAW99408	Staphylokinase Sak
562	6	4.4	1189	21	AAAG2348	Arabidopsis thalia	635	5	3.7	17	22	AAW99409	Staphylokinase Sak
563	6	4.4	1244	19	AAW59358	Human retinal dege	636	5	3.7	17	22	AAW898090	Chimeric L6-VH plu
564	6	4.4	1270	22	AAE01982	Human ATPase-relat	637	5	3.7	18	21	AAW828949	Peptide Fel 34.. F
565	6	4.4	1276	21	AAAG31886	Arabidopsis thalia	638	5	3.7	18	21	AAW87690	Feline human TRFP
566	6	4.4	1285	19	AAW47000	HIS tagged thioered	639	5	3.7	18	21	AAW90120	Cat TRFP derived p
567	6	4.4	1294	20	AAAY39633	Zea mays Rpg1-2 pr	640	5	3.7	18	21	AAW51487	Human TRFP derived
568	6	4.4	1336	21	AAAG48590	Arabidopsis thalia	641	5	3.7	18	21	AAW59134	Yeast STE20 Gbd/CR
569	6	4.4	1354	18	AAW23654	Physiologically ac	642	5	3.7	18	22	AAW20421	Anti-Fix/FTXa anti
570	6	4.4	1354	19	AAW71020	A modified Rho tar	643	5	3.7	19	13	AAW26956	Human T lymphocyte
571	6	4.4	1354	20	AAAY07082	Renal cancer assoc	644	5	3.7	19	16	AAW80836	Human CD40 ligand
572	6	4.4	1405	19	AAW56101	Enhanced green flu	645	5	3.7	19	17	AAW49182	Human leucocyte an
573	6	4.4	1426	22	AAE01984	Human ATPase-relat	646	5	3.7	19	18	AAW16904	Helicobacter pylor
574	6	4.4	1426	22	AAE01984	Human ATPase-relat	647	5	3.7	19	18	AAW16905	Helicobacter pylor
575	6	4.4	1957	18	AAAB61995	Rat peripheral ner	648	5	3.7	19	20	AAW78360	MAM mitogenic acti
576	6	4.4	1957	18	AAW21740	Variant rat DRG (S	649	5	3.7	19	21	AAAY0296	Plasmodium falcipa
577	6	4.4	2132	18	AAW21737	Wild type rat DRG	650	5	3.7	19	22	AAW5085	Peptide #1519 enco
578	6	4.4	2132	18	AAW21739	Variant rat DRG (S	651	5	3.7	19	22	AAW27533	Peptide #1570 enco
579	6	4.4	2206	21	AAW18254	Plasmodium falcipa	652	5	3.7	19	22	AAW02821	Peptide #1503 enco
580	6	4.4	2485	21	AAW18172	Plasmodium falcipa	653	5	3.7	19	22	AAW30690	Rhodococcus haloa
581	6	4.4	2841	21	AAW07581	Protein encoded by	654	5	3.7	20	19	AAW76178	Feline FLAP40 pro
582	6	4.4	3079	15	AAW59926	GAP protein Irax2.	655	5	3.7	20	21	AAW45069	D. immitis aromati
583	6	4.4	3519	21	AAAY92709	S. anthiolicus 8,	656	5	3.7	20	22	AAW868950	Arabidopsis thalia
584	6	4.4	4150	21	AAAY92707	S. anthiolicus 8,	657	5	3.7	21	15	AAW46969	J alpha sequence (
585	6	4.4	5701	22	AAW05328	Peptide #4010 enco	658	5	3.7	21	18	AAW35466	Immunogenic agent
586	6	4.4	26926	22	AAW05396	Human titin (cone	659	5	3.7	21	21	AAW35466	Peptide encoded by
587	5	3.7	5	19	AAW51489	Peptide epitope 15	660	5	3.7	21	21	AAW87709	Feline human TRFP
588	5	3.7	6	19	AAW51524	Pre-sequence which	661	5	3.7	21	21	AAAY90135	TRFP derived pepit
589	5	3.7	6	19	AAW51526	Pre-sequence which	662	5	3.7	21	21	AAAY51506	Human TRFP derived
590	5	3.7	6	19	AAW51536	Pre-sequence which	663	5	3.7	22	16	AAW83834	Silkworm phenoloxi
591	5	3.7	6	19	AAW51537	Pre-sequence which	664	5	3.7	22	20	AAW19196	Lecithin:cholester
592	5	3.7	7	17	AAW45651	GL-NAc acylase al	665	5	3.7	22	20	AAW18942	Lecithin:cholester
593	5	3.7	7	18	AAW36924	Peptide epitope 19	666	5	3.7	22	20	AAW18688	Lecithin:cholester
594	5	3.7	7	18	AAW06506	DPPIV peptide frag	667	5	3.7	22	20	AAW18425	Lecithin:cholester
595	5	3.7	7	22	AAW85436	Serum DPPIV trypti	668	5	3.7	22	20	AAW83093	Peptide from the 3

669	5	3.7	22	21	AAV77782	Core polypeptide u	742	5	3.7	27	22	AAAI4219	Peptide #653 encod
670	5	3.7	22	22	AAAI1980	Peptide #6017 enco	743	5	3.7	27	22	AAAI20818	Peptide #7252 enco
671	5	3.7	22	22	AAAB78467	Core polypeptide #	744	5	3.7	27	22	AAAI26629	Peptide #666 enco
672	5	3.7	23	15	AAAB6968	Core polypeptide #	745	5	3.7	27	22	AAAI36274	Peptide #10311 enc
673	5	3.7	23	19	AAAB58706	Cytotoxic lymphocy	746	5	3.7	27	22	AAAI01951	Peptide #633 encod
674	5	3.7	23	21	AAV96435	Mutant human TGF-b	747	5	3.7	28	19	AAAI46283	Moraxella catarrha
675	5	3.7	23	21	AAV83628	Peptide fragment o	748	5	3.7	29	20	AAV29691	Moraxella catarrha
676	5	3.7	23	21	AAV77779	Core polypeptide u	749	5	3.7	29	21	AAV51912	Basic peptide NST3
677	5	3.7	23	21	AAV77780	Core polypeptide u	750	5	3.7	30	21	AAV83795	Ationic-phospholip
678	5	3.7	23	21	AAV77781	Core polypeptide u	751	5	3.7	30	19	AAV77620	Staphylococcus aur
679	5	3.7	23	21	AAV77783	Core polypeptide u	752	5	3.7	31	21	AAV83798	Ationic-phospholip
680	5	3.7	23	21	AAV77784	Core polypeptide u	753	5	3.7	32	9	AAAB80106	Encoded by Staphyl
681	5	3.7	23	21	AAV89820	Core polypeptide f	754	5	3.7	32	21	AAAB38146	Human secreted pro
682	5	3.7	23	21	AAV89821	Core polypeptide f	755	5	3.7	33	17	AAAB95174	Stem cell factor C
683	5	3.7	23	21	AAV65740	Breast cancer susc	756	5	3.7	33	21	AAAB16943	Bacteriophage Dp-1
684	5	3.7	23	22	AAAB78221	Core polypeptide T	757	5	3.7	33	21	AAAB89721	Core polypeptide f
685	5	3.7	23	22	AAAB78222	Core polypeptide T	758	5	3.7	33	21	AAAB89722	Core polypeptide f
686	5	3.7	23	22	AAAB78464	Core polypeptide #	759	5	3.7	33	22	AAAB78121	Core polypeptide T
687	5	3.7	23	22	AAAB78465	Core polypeptide #	760	5	3.7	33	22	AAAB78122	Core polypeptide T
688	5	3.7	23	22	AAAB78466	Core polypeptide #	761	5	3.7	34	21	AAV83796	Ationic-phospholip
689	5	3.7	23	22	AAAB78468	Core polypeptide #	762	5	3.7	34	22	AAAI18944	Peptide #5378 enco
690	5	3.7	23	22	AAAB78469	Core polypeptide #	763	5	3.7	34	22	AAAI28826	Peptide #2863 enco
691	5	3.7	24	12	AAAI3011	Human Factor VIII (	764	5	3.7	34	22	AAAI31520	Peptide #5557 enco
692	5	3.7	24	21	AAV65741	Breast cancer susc	765	5	3.7	34	22	AAU05883	Cone snail O-supe
693	5	3.7	24	22	AAAI6581	Peptide #6115 enco	766	5	3.7	34	22	AAAI04069	Peptide #2751 enco
694	5	3.7	24	22	AAAI6581	Peptide #6947 enco	767	5	3.7	35	14	AAAB45158	Boltoectin one-cha
695	5	3.7	25	19	AAAI32910	Peptide #6425 enco	768	5	3.7	35	21	AAAB44354	Human secreted pro
696	5	3.7	25	22	AAAI19991	Peptide #6425 enco	769	5	3.7	35	21	AAV886649	Hybrid polypeptide
697	5	3.7	25	22	AAAI3693	Peptide #7730 enco	770	5	3.7	35	21	AAV89717	Core polypeptide f
698	5	3.7	26	14	AAAB36544	Peptide 2, Felis,	771	5	3.7	35	21	AAV89730	Core polypeptide f
699	5	3.7	26	14	AAAB41977	Human T cell react	772	5	3.7	35	21	AAV89774	Core polypeptide f
700	5	3.7	26	15	AAAB52478	Human heavy chain	773	5	3.7	35	21	AAV89792	Core polypeptide f
701	5	3.7	26	15	AAAB52480	Human heavy chain	774	5	3.7	35	21	AAV89799	Core polypeptide f
702	5	3.7	26	15	AAAB52462	Human heavy chain	775	5	3.7	35	21	AAV89800	Core polypeptide f
703	5	3.7	26	15	AAAB52463	Human heavy chain	776	5	3.7	35	22	AAAB68863	Core polypeptide #
704	5	3.7	26	15	AAAB52466	Human heavy chain	777	5	3.7	35	22	AAAB78118	Core polypeptide T
705	5	3.7	26	15	AAAB52474	Human heavy chain	778	5	3.7	35	22	AAAB78131	Core polypeptide T
706	5	3.7	26	15	AAAB52475	Human heavy chain	779	5	3.7	35	22	AAAB78132	Core polypeptide T
707	5	3.7	26	15	AAAB52476	Human heavy chain	780	5	3.7	35	22	AAAB78193	Core polypeptide T
708	5	3.7	26	15	AAAB52453	Human heavy chain	781	5	3.7	35	22	AAAB78200	Core polypeptide T
709	5	3.7	26	15	AAAB52455	Human heavy chain	782	5	3.7	35	22	AAAB78201	Core polypeptide T
710	5	3.7	26	15	AAAB52456	Human heavy chain	783	5	3.7	36	12	AAAI2372	Antihaemostatic pe
711	5	3.7	26	15	AAAB52458	Human heavy chain	784	5	3.7	36	20	AAV35985	Extended human sec
712	5	3.7	26	15	AAAB52489	Mouse heavy chain	785	5	3.7	36	21	AAV88688	Core polypeptide f
713	5	3.7	26	15	AAAB52290	Mouse heavy chain	786	5	3.7	36	21	AAV89689	Core polypeptide f
714	5	3.7	26	15	AAAB52298	Mouse heavy chain	787	5	3.7	36	22	AAAI4672	Peptide #1106 enco
715	5	3.7	26	15	AAAB52273	Mouse heavy chain	788	5	3.7	36	22	AAAI27091	Peptide #1128 enco
716	5	3.7	26	15	AAAB52276	Mouse heavy chain	789	5	3.7	36	22	AAAI02395	Peptide #1077 enco
717	5	3.7	26	15	AAAB52277	Mouse heavy chain	790	5	3.7	36	22	AAAB20035	Zinc finger helica
718	5	3.7	26	15	AAAB52278	Mouse heavy chain	791	5	3.7	36	22	AAAB78089	Core polypeptide T
719	5	3.7	26	15	AAAB52279	Mouse heavy chain	792	5	3.7	36	22	AAAB78090	Core polypeptide T
720	5	3.7	26	15	AAAB52281	Mouse heavy chain	793	5	3.7	36	22	AAAB59477	Human secreted pro
721	5	3.7	26	15	AAAB52282	Mouse heavy chain	794	5	3.7	37	20	AAAB88395	Human zneul epitop
722	5	3.7	26	15	AAAB52283	Mouse heavy chain	795	5	3.7	37	21	AAAB63167	Human secreted pro
723	5	3.7	26	15	AAAB52260	Mouse heavy chain	796	5	3.7	37	21	AAAB45268	Human secreted pro
724	5	3.7	26	15	AAAB52261	Mouse heavy chain	797	5	3.7	37	21	AAV89732	Core polypeptide f
725	5	3.7	26	15	AAAB52262	Mouse heavy chain	798	5	3.7	37	22	AAAI15945	Peptide #6383 enco
726	5	3.7	26	15	AAAB52263	Mouse heavy chain	799	5	3.7	37	22	AAAB3601	Peptide #7638 enco
727	5	3.7	26	15	AAAB52264	Mouse heavy chain	800	5	3.7	37	22	AAAB78133	Core polypeptide T
728	5	3.7	26	15	AAAB52266	Mouse heavy chain	801	5	3.7	37	22	AAAB65949	Human secreted pro
729	5	3.7	26	15	AAAB52267	Mouse heavy chain	802	5	3.7	38	21	AAV89765	Core polypeptide f
730	5	3.7	26	15	AAAB52268	Mouse heavy chain	803	5	3.7	39	20	AAV29822	Ahesion epitope N
731	5	3.7	26	15	AAAB52269	Mouse heavy chain	804	5	3.7	39	21	AAAB57001	Human prostate can
732	5	3.7	26	15	AAAB52270	Mouse heavy chain	805	5	3.7	39	21	AAV89731	Core polypeptide f
733	5	3.7	26	19	AAAB6253	Moraxella catarrha	806	5	3.7	39	21	AAV89744	Core polypeptide f
734	5	3.7	26	19	AAAB6256	Moraxella catarrha	807	5	3.7	39	21	AAV89745	Core polypeptide f
735	5	3.7	26	21	AAAB28943	Peptide 2 derived	808	5	3.7	39	21	AAV89748	Core polypeptide f
736	5	3.7	26	21	AAAB18497	Peptide substrate	809	5	3.7	39	21	AAV89766	Core polypeptide f
737	5	3.7	26	21	AAV87684	Feline human TRFP	810	5	3.7	39	22	AAAB78132	Core polypeptide T
738	5	3.7	26	21	AAV90109	Cat TRFP derived p	811	5	3.7	39	22	AAAB78145	Core polypeptide T
739	5	3.7	26	21	AAV83797	Antonic-phospholip	812	5	3.7	39	22	AAAB78146	Core polypeptide T
740	5	3.7	26	21	AAV51481	Human TRFP derived	813	5	3.7	39	22	AAAB78149	Core polypeptide T
741	5	3.7	27	19	AAV86016	S. pneumoniae deri	814	5	3.7	39	22	AAAB78166	Core polypeptide T



815	5	3.7	39	22	AAB78167	Core polypeptide T	888	5	3.7	55	20	AAV02669	Human secreted pro
816	5	3.7	40	17	AAR96936	Thymosin beta 14.	889	5	3.7	55	21	AAB34671	Gene 35 human secr
817	5	3.7	40	21	AAB38431	Fragment of human	890	5	3.7	55	21	AAB34672	Human secreted pro
818	5	3.7	40	21	AAB12944	Neurotransmission	891	5	3.7	55	21	AAAG24546	Arabidopsis thalia
819	5	3.7	40	21	AAV80271	Thymosin beta 4 pe	892	5	3.7	56	22	AAAM18283	Peptide #4717 enco
820	5	3.7	40	21	AAV89747	Core polypeptide f	893	5	3.7	56	22	AAAM30773	Peptide #4810 enco
821	5	3.7	40	22	AAB78148	Core polypeptide T	894	5	3.7	56	22	AAAM36305	Peptide #10342 enco
822	5	3.7	41	19	AAW53674	FIV ppr clone 34 O	895	5	3.7	56	22	AAAM05896	Peptide #4578 enco
823	5	3.7	41	20	AAV12877	Human 5' EST secre	896	5	3.7	56	22	AAAG73383	Human gene 2-encod
824	5	3.7	41	21	AAV89749	Core polypeptide f	897	5	3.7	57	22	AAB81179	Chimeric transglut
825	5	3.7	41	22	AAAM15285	Peptide #1719 enco	898	5	3.7	57	21	AAAG14328	Arabidopsis thalia
826	5	3.7	41	22	AAU065521	Human CYP2B6 prote	899	5	3.7	57	21	AAAG54611	Zea mays protein f
827	5	3.7	41	22	AAU065522	Human CYP2B6 prote	900	5	3.7	57	22	AAAM15951	Peptide #2385 enco
828	5	3.7	41	22	AAB81172	Transglutaminase r	901	5	3.7	57	22	AAAM20678	Peptide #7112 enco
829	5	3.7	41	22	AAB81177	Modified transglut	902	5	3.7	57	22	AAAM28454	Peptide #2491 enco
830	5	3.7	41	22	AAB81178	Modified transglut	903	5	3.7	57	22	AAAM34808	Peptide #8445 enco
831	5	3.7	41	22	AAB78150	Core polypeptide T	904	5	3.7	57	22	AAAM35838	Peptide #9875 enco
832	5	3.7	42	21	AAV89746	Core polypeptide f	905	5	3.7	57	22	AAAM36456	Peptide #10493 enc
833	5	3.7	42	22	AAAM14962	Peptide #1396 enco	906	5	3.7	57	22	AAAM03689	Peptide #2371 enco
834	5	3.7	42	22	AAAM27396	Peptide #1433 enco	907	5	3.7	58	17	AAW34239	SH3 domain of huma
835	5	3.7	42	22	AAAM02688	Peptide #1370 enco	908	5	3.7	58	19	AAAB38588	Streptococcus pneu
836	5	3.7	42	22	AAB78147	Core polypeptide T	909	5	3.7	58	21	AAAB53983	Human colon cancer
837	5	3.7	43	16	AAR70973	Core polypeptide f	910	5	3.7	58	21	AAAG09957	Arabidopsis thalia
838	5	3.7	43	17	AAV95471	HARDS virus peptid	911	5	3.7	58	21	AAAG40074	Arabidopsis thalia
839	5	3.7	43	19	AAW69984	v44, monoclonal an	912	5	3.7	58	21	AAAG59902	Arabidopsis thalia
840	5	3.7	43	21	AAV86370	Yeast tropomyosin	913	5	3.7	58	21	AAAG03445	Human secreted pro
841	5	3.7	43	22	AAAM33104	Human gene 11-enco	914	5	3.7	58	22	AAAG03445	Human EST encoded
842	5	3.7	43	22	AAAG74160	Peptide #7141 enco	915	5	3.7	58	22	AAE01280	Human gene 11 enco
843	5	3.7	43	22	AAB81171	Human colon cancer	916	5	3.7	59	22	AAAM17863	Peptide #4297 enco
844	5	3.7	44	16	AAAR83836	Transglutaminase r	917	5	3.7	59	22	AAAM30370	Peptide #4407 enco
845	5	3.7	44	16	AAV74023	Silkworm prophenol	918	5	3.7	59	22	AAAB34380	Gene 14 human secr
846	5	3.7	44	21	AAAB43473	Human prostate tum	919	5	3.7	60	21	AAAB34381	Human secreted pro
847	5	3.7	44	21	AAV80272	Human Cancer assoc	920	5	3.7	60	21	AAAG16411	Arabidopsis thalia
848	5	3.7	44	22	AAB81173	Thymosin beta 4 pe	921	5	3.7	60	22	AAAM15955	Peptide #2389 enco
849	5	3.7	45	16	AAAR74269	Transglutaminase r	922	5	3.7	60	22	AAAM17151	Peptide #3585 enco
850	5	3.7	45	18	AAW36056	Borrelia OspC SSAL	923	5	3.7	60	22	AAAM28457	Peptide #2494 enco
851	5	3.7	45	18	AAAM14281	Human thymosin bet	924	5	3.7	60	22	AAAM29134	Peptide #3171 enco
852	5	3.7	45	19	AAW68573	Human neuroblastom	925	5	3.7	60	22	AAAM29643	Peptide #3680 enco
853	5	3.7	45	19	AAW72683	Rat thymosin-beta1	926	5	3.7	60	22	AAAM03693	Peptide #2375 enco
854	5	3.7	45	19	AAW46486	Ikari protein SEQ	927	5	3.7	60	22	AAAM04366	Peptide #3048 enco
855	5	3.7	45	19	AAW44275	Human thymosin bet	928	5	3.7	60	22	AAAM04846	Peptide #3528 enco
856	5	3.7	45	20	AAV33663	Human thymosin bet	929	5	3.7	61	15	AAAR66849	Rice mitochondrial
857	5	3.7	45	20	AAV13083	Streptoveritciliu	930	5	3.7	61	15	AAAR74751	Rice mitochondrial
858	5	3.7	45	21	AAAG06157	Human secreted pro	931	5	3.7	62	14	AAAR38495	B. nodosus pillin P
859	5	3.7	45	21	AAV76812	Arabidopsis thalia	932	5	3.7	62	22	AAAM19687	Peptide #6121 enco
860	5	3.7	45	22	AAAM14966	Human thymosin bet	933	5	3.7	62	22	AAAM22046	Peptide #8480 enco
861	5	3.7	45	22	AAAM27400	Peptide #1400 enco	934	5	3.7	62	22	AAAM32928	Peptide #6965 enco
862	5	3.7	45	22	AAB81159	Peptide #1437 enco	935	5	3.7	62	22	AAAM38427	Peptide #12464 enc
863	5	3.7	45	22	AAB81175	Pre-pro transgluta	936	5	3.7	63	22	AAAM14318	Peptide #752 encod
864	5	3.7	45	22	AAB81176	Modified transglut	937	5	3.7	63	22	AAAM26728	Peptide #765 encod
865	5	3.7	46	12	AAAR10755	Non-A non-B hepati	938	5	3.7	63	22	AAAM02047	Peptide #729 encod
866	5	3.7	47	22	AAAM21200	Peptide #7634 enco	939	5	3.7	63	22	AAU02804	Taxus cuspidata RT
867	5	3.7	47	22	AAAM37417	Peptide #11454 enco	940	5	3.7	64	22	AAAM15723	Peptide #2157 enco
868	5	3.7	48	22	AAAM35360	Peptide #9397 enco	941	5	3.7	64	22	AAAM28228	Peptide #2265 enco
869	5	3.7	49	21	AAAG20089	Arabidopsis thalia	942	5	3.7	64	22	AAAM31425	Peptide #5462 enco
870	5	3.7	50	20	AAV12009	Human 5' EST secre	943	5	3.7	64	22	AAAM03457	Peptide #2139 enco
871	5	3.7	50	21	AAB39098	Human secreted pro	944	5	3.7	64	22	AAE01816	Human gene 10 enco
872	5	3.7	50	21	AAAG56198	Arabidopsis thalia	945	5	3.7	65	21	AAAG55992	Arabidopsis thalia
873	5	3.7	50	22	AAAM35311	Peptide #9348 enco	946	5	3.7	65	22	AAAM15961	Peptide #2395 enco
874	5	3.7	51	22	AAAM20000	Peptide #6434 enco	947	5	3.7	65	22	AAAM28463	Peptide #2500 enco
875	5	3.7	51	22	AAAM33719	Peptide #7736 enco	948	5	3.7	65	22	AAAM03698	Peptide #2380 enco
876	5	3.7	51	22	AAAM34869	Peptide #8906 enco	949	5	3.7	65	22	AAAG81287	Human AFP protein
877	5	3.7	52	20	AAV60333	Human normal blad	950	5	3.7	66	21	AAAB34637	Gene 3 human secre
878	5	3.7	52	21	AAAG46287	Arabidopsis thalia	951	5	3.7	66	22	AAAB34638	Human secreted pro
879	5	3.7	52	21	AAAG57145	Arabidopsis thalia	952	5	3.7	66	22	AAAM20089	Peptide #6523 enco
880	5	3.7	52	22	AAAM37794	Peptide #11831 enc	953	5	3.7	66	22	AAAM33953	Peptide #7990 enco
881	5	3.7	52	22	AAAG75231	Human colon cancer	954	5	3.7	67	14	AAAR38496	B. nodosus pillin P
882	5	3.7	53	21	AAAG58688	Arabidopsis thalia	955	5	3.7	67	14	AAAG59713	Arabidopsis thalia
883	5	3.7	53	22	AAAM34088	Peptide #8125 enco	956	5	3.7	67	21	AAAG59816	Arabidopsis thalia
884	5	3.7	53	22	AAAM34627	Peptide #8664 enco	957	5	3.7	67	22	AAAM17275	Peptide #3709 enco
885	5	3.7	54	20	AAW78358	M. arthritis mit	958	5	3.7	67	22	AAAM29773	Peptide #3810 enco
886	5	3.7	54	21	AAV54438	Peptide which mimi	959	5	3.7	67	22	AAAM04963	Peptide #3645 enco
887	5	3.7	54	22	AAAM36386	Peptide #10423 enc	960	5	3.7	68	21	AAAG00889	Human secreted pro



Db	1	ssfdkgkykkggdaasyfepgpylmnnvrgvdsknqellspbyvteipkpgtlltckek1	60
Qy	61	EYYEWALDAFAKKEFRVVELDPSAKLEVTYYDKNKKKEETKSPITEKGFVPDISEHT	120
Db	61	EYYEWALDAFAKKEFRVVELDPSAKLEVTYYDKNKKKEETKSPITEKGFVPDISEHT	120
Qy	121	KNPGFNLTAKVIEKK	136
Db	121	KNPGFNLTAKVIEKK	136
RESULT 2			
AAW03086	standard; protein: 136 AA.		
AAW03086;			
19-FEB-1997	(first entry)		
DE	Wild-type staphylokinase.		
KW	Staphylokinase; mutant; mutein; variant; immunogenicity; decrease;		
KW	derivative; SakSTAR; arterial thrombosis; thrombolytic agent.		
OS	Staphylococcus aureus.		
Key	Location/Qualifiers		
Region	5..6	/label= M20_epitope	
Region	8..10	/label= M21_epitope	
Region	11..14	/label= M1_epitope	
Region	19	/label= M22_epitope	
Region	33..35	/label= M2_epitope	
Region	35..38	/label= M3_epitope	
Region	46..50	/label= M4_epitope	
Region	57..59	/label= M5_epitope	
Region	61..65	/label= M6_epitope	
Region	65..69	/label= M7_epitope	
Region	74..77	/label= M8_epitope	
Region	80..82	/label= M9_epitope	
Region	86..88	/label= M10_epitope	
Region	93..94	/label= M11_epitope	
Region	96..98	/label= M12_epitope	
Region	99..100	/label= M13_epitope	
Region	99..102	/label= M14_epitope	
Region	108..109	/label= M15_epitope	
Region	115..119	/label= M16_epitope	
Region	119..121	/label= M17_epitope	
Region	130	/label= M18_epitope	
Region	134..136	/label= M19_epitope	
PN	W09621016-A2.		

[illegible]

```

PF 01-DEC-2000; 2000WO-EP12229.
XX
XX 02-DEC-1999; 99EP-0204093.
XX
XX (THRO-) THROMB-X NV.
XX
XX Warmerdam PAM, Plaaisance SDNGH, Collen DJ, De Maeyer MCH;
XX
XX WPI: 2001-374786/39.
XX
XX Reducing immunogenicity of protein (P) by eliminating T cell epitopes
XX in test peptides having amino acid sequence corresponding to (P) and
XX modifying amino acid sequence of (P) according to test peptide
XX modifications -
XX
XX Disclosure; Fig 1: 50pp; English.
XX
XX The present invention describes a method for reducing the immunogenicity
XX of a peptide or protein (I), where (I) can be staphylokinase (SakSTAR)
XX isolated from Staphylococcus aureus. The method involves designing a
XX series of overlapping test peptides having an amino acid sequence
XX corresponding to (I), modifying test peptides which are identified to
XX comprise one or more T-cell epitopes, such that they are reduced or
XX eliminated, and repeating the T-cell eliminating modifications for (I)
XX to produce a modified peptide or protein. Methods from the present
XX invention can be used for the treatment, diagnosis or prophylaxis or
XX for the preparation of a pharmaceutical composition for the treatment,
XX diagnosis or prophylaxis of a human subject. Staphylokinase is used as a
XX potent thrombolytic agent in patients with acute myocardial infarction.
XX The method is useful for reducing cell based immunogenicity of non-human
XX proteins such as streptokinase or antibodies or their fragments, from
XX other species, for diagnostics and treatment of human disease. AAB99400
XX to AAB99449 represent amino acid sequences used in the exemplification
XX of the present invention.
XX
XX Sequence 136 AA:
SQ

```

Query Match 100.0%; Score 136; DB 22; Length 136;  
Best Local Similarity 100.0%; Pred. No. 3.9e-131;  
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 SSFSDGKTKKKGDASYFEPTGTYLMVNTGVDSKGNELSPHYVEPIKPGTTLRKEKI 60
   |||||||
DB 1 sssfdkgkykkgdasyfepctgpylmvntgvyskgnellsphyvefpikpgttlkeki 60
QY 61 EYVEWALDATAYKEFRVVELDPSAKIEVYVYDKNKKKEETKSFPTTEKGFVVPDLSEHI 120
   |||||||
DB 61 eyvewaldataykefrvveldpsakievtydydknkkkeetsfpktegfvvpdlsehi 120
QY 121 KNGFNLITKRVYIEKK 136
   |||||||
DB 121 knpgfnlilkvylekk 136

```

RESULT 4  
AAB61909 standard; Protein: 136 AA.  
XX  
XX AAB61909;  
XX  
XX 08-MAY-2001 (first entry)  
XX  
XX S. aureus staphylokinase G34S variant.  
XX  
XX Immunogenicity; staphylokinase; variant; stability.  
XX  
XX Staphylococcus aureus.  
XX  
XX Key Location/Qualifiers  
XX FH Misc-difference 34 /label- G34S  
XX FT /note- "wild-type Gly is replaced by Ser"  
XX FT

```

XX
XX WO200104287-A1.
XX
XX 18-JAN-2001.
XX
XX 06-JUL-2000; 2000WO-DK00371.
XX
XX 07-JUL-1999; 99DK-0000988.
XX
XX 27-AUG-1999; 99DK-0001196.
XX
XX 02-MAR-2000; 2000DK-0000339.
XX
XX 18-MAY-2000; 2000DK-0000804.
XX
XX (MAXY-) MAXYGEN APS.
XX
XX Halkier T, Pedersen AH, Okkels JS;
XX
XX WPI: 2001-138342/14.
XX
XX Producing polypeptides with altered immunogenicity or improved
XX stability, comprises expressing a diversified nucleotide sequence
XX population and selecting polypeptides with altered immunogenicity or
XX improved stability -
XX
XX Example 1; Page 76; 83pp; English.
XX
XX The invention relates to a method of altering immunogenicity and/or
XX increasing stability of a polypeptide of interest. The method comprises
XX (a) expressing a diversified population of nucleotide sequences encoding
XX a polypeptide of interest; (b) screening the polypeptides expressed for
XX function, immunogenicity and/or stability; and (c) selecting functional
XX polypeptides with altered immunogenicity and/or increased stability. The
XX method is used to improve the properties of polypeptides, in particular
XX to alter the immunogenicity and/or increase the functional in vivo half-
XX life of the polypeptide. The method uses a high throughput system that
XX makes it possible to search several orders of magnitude more polypeptides
XX than is possible by previously known approaches. This enhances the chance
XX of finding the optimal variant from the many thousands of variants that
XX CC may be produced. The present sequence represents S. aureus staphylokinase
XX variant sequence, used to exemplify the method of the invention.
XX
XX Sequence 136 AA:
SQ

```

Query Match 100.0%; Score 136; DB 22; Length 136;  
Best Local Similarity 100.0%; Pred. No. 3.9e-131;  
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 SSFSDGKTKKKGDASYFEPTGTYLMVNTGVDSKGNELSPHYVEPIKPGTTLRKEKI 60
   |||||||
DB 1 sssfdkgkykkgdasyfepctgpylmvntgvyskgnellsphyvefpikpgttlkeki 60
QY 61 EYVEWALDATAYKEFRVVELDPSAKIEVYVYDKNKKKEETKSFPTTEKGFVVPDLSEHI 120
   |||||||
DB 61 eyvewaldataykefrvveldpsakievtydydknkkkeetsfpktegfvvpdlsehi 120
QY 121 KNGFNLITKRVYIEKK 136
   |||||||
DB 121 knpgfnlilkvylekk 136

```

RESULT 5  
AAR39150 standard; Protein: 137 AA.  
XX  
XX AAR39150;  
XX  
XX 03-DEC-1993 (first entry)  
XX  
XX Staphylokinase SAK-STAR.  
XX  
XX Signal sequence, plasminogen activator; thrombosis; staphylokinase;  
XX SAK.  
XX

OS Staphylococcus aureus strain 23.  
XX  
PN W09313209-A.  
XX  
PD 08-JUL-1993.  
XX  
PE 28-DEC-1992; 92WO-EP02989.  
XX  
PR 30-DEC-1991; 91DE-4143279.  
PR 22-JUN-1992; 92DE-4220516.  
PR 01-DEC-1992; 92DE-4240801.  
XX  
PA (MEDA-) MEDAC GES KLINISCHE SPEZIALPRAPARATE.  
XX  
PI Albrecht S, Behnke D, Guehrs K, Hartmann M, Schloft B;  
XX  
DR WPI; 1993-227325/28.  
DR N-PSDB; AAO44270.  
XX  
PT Staphylo-kinase (SAK) sequences lacking sequences for signal  
PT peptide(s) - for prodn. of proteins used as plasminogen  
PT activators in thrombosis treatment, and monoclonal antibodies  
PT against SAK  
XX  
PS Claim 3; Fig 3; 99pp; German.  
XX  
CC DNA encoding SAK lacking the signal peptide, is expressed  
CC intracellularly. This avoids the problem of fast degradation of the  
CC polypeptides or destruction of the host when expressed into the  
CC medium or into the periplasm respectively. High expression is  
CC possible and the chemically induced overprodn. is easy to handle.  
CC Also, the prods. are homogeneous.  
CC SAK-polypeptide derivs. are plasminogen activators for the  
CC treatment of thrombosis.  
XX  
SQ Sequence 137 AA:  
  
Query Match 100.0%; Score 136; DB 14; Length 137;  
Best Local Similarity 100.0%; Pred. No. 3.9e-131;  
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 SSSFDKGGKRRGDDASYFEPTGPIYLMVNTGVDSKGNELLSPHYVEFPKPGTTLTKKXI 60  
DB 2 sssfdkgykkgddasyfepitgpylmvntgydskgnellsplyvefpikpgtlltkexi 61  
OY 61 EYVEMALDATAYKERYVELDPSAKIEVTYYDKNKKKEETSFPITTEGFWVPDLSEHI 120  
DB 62 eyvewaldataykefrvvelpsakievtyydknkkkeetsfpitkxgfvvpdlsehi 121  
OY 121 KNPGFNLITRKVI 136  
DB 122 knpgfnlilkvilekk 137  
  
RESULT 6  
AAW03102  
ID AAW03102 standard; protein; 136 AA.  
XX  
AC AAW03102;  
XX  
DT 19-FEB-1997 (first entry)  
XX  
DE Staphylokinase derivative M19 with altered charge cluster 19.  
XX  
KW Staphylokinase; mutant; mutein; variant; immunogenicity; decrease;  
KW derivative; SAKSTAR; arterial thrombosis; thrombolytic agent.  
XX  
OS Staphylococcus aureus.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Region 134..136

FT /label= altered\_M19\_cluster  
FT /note= "wild-type cluster 19 sequence Glu-Lys-Lys  
FT has been replaced by Ala-Ala-Ala to reduce  
FT immunogenicity"  
XX  
XX W09621016-A2.  
XX  
PD 11-JUL-1996.  
XX  
PE 03-JAN-1996; 96WO-EP00081.  
XX  
PR 17-NOV-1995; 95JP-0299781.  
PR 06-JAN-1995; 95EP-0200023.  
PR 11-JAN-1995; 95OS-0371505.  
PR 09-JUN-1995; 95EP-0201531.  
PR 06-JUL-1995; 95US-0499092.  
XX  
PA (COLL/) COLLEN D.  
PA (LEUV-) LEUVEN RES & DEV VZW.  
XX  
PI Collen D;  
XX  
DR WPI; 1996-333991/33.  
XX  
PT New staphylokinase derivs. having reduced immunogenicity - useful  
PT for treating arterial thrombosis  
XX  
PS Example 2; Page -: 58pp; English.  
XX  
CC Staphylokinase derivatives showing a reduced immunogenicity as  
CC compared to wild-type staphylokinase are claimed. The derivatives  
CC are useful as thrombolytic agents to treat arterial thrombosis and  
CC are pref. produced by eliminating at least one of the epitopes  
CC indicated in the features table. The epitopes are destroyed by  
CC replacing one or more amino acid residues in a charge cluster by an  
CC Ala residue. Mutations are introduced using site-directed  
CC mutagenesis on wild-type staphylokinase genomic DNA from the S.aureus  
CC lysogenic strain SAKSTAR. The present sequence is a specific  
CC example of a staphylokinase derivative according to the invention;  
CC the sequence does not appear in the specification but can be generated  
CC from the wild-type staphylokinase sequence using the description given  
CC in Table 1 of the disclosure.  
XX  
SQ Sequence 136 AA:  
  
Query Match 97.8%; Score 133; DB 17; Length 136;  
Best Local Similarity 100.0%; Pred. No. 4.5e-128;  
Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 SSSFDKGGKRRGDDASYFEPTGPIYLMVNTGVDSKGNELLSPHYVEFPKPGTTLTKKXI 60  
DB 1 sssfdkgykkgddasyfepitgpylmvntgydskgnellsplyvefpikpgtlltkexi 60  
OY 61 EYVEMALDATAYKERYVELDPSAKIEVTYYDKNKKKEETSFPITTEGFWVPDLSEHI 120  
DB 61 eyvewaldataykefrvvelpsakievtyydknkkkeetsfpitkxgfvvpdlsehi 120  
OY 121 KNPGFNLITRKVI 133  
DB 122 knpgfnlilkvilekk 133  
  
RESULT 7  
AAW03085  
ID AAW03085 standard; protein; 136 AA.  
XX  
AC AAW03085;  
XX  
DT 19-FEB-1997 (first entry)  
XX  
DE Staphylokinase derivative M19 with altered charge cluster 19.  
XX



```

RESULT 9
AAV15025
ID AAV15025 standard; Protein: 136 AA.
XX
XX
AC AAV15025;
XX
XX
DT 03-NOV-1999 (first entry)
XX
XX
DE Staphylokinase variant SakSTAR (V132P).
XX
XX
KW Staphylokinase; variant; immunogenic; specificity; derivative; cysteine;
KW SakSTAR (V132P); anti-thrombotic; fibrinolytic; cardiant; veterinary;
KW Staphylococcus aureus; myocardial infarction; arterial thrombosis;
KW monoclonal antibody; polyethylene glycol; PEG; plasma clearance.
XX
OS Staphylococcus aureus.
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 132 /note="Wild type Val is substituted by Thr"
FT
XX
XX
PN MO9940198-A2.
XX
XX
PD 12-AUG-1999.
XX
XX
PE 04-FEB-1999; 99WO-EP00748.
XX
XX
PR 06-FEB-1998; 98EP-0200365.
PR 04-FEB-1998; 98EP-0200323.
XX
XX
PA (COLL/) COLLEN D J.
PA (LEUV-) LEUVEN RES & DEV VZW.
XX
XX
PI Collen DJ;
XX
XX
DR WPI: 1999-508504/42.
XX
XX
PT Staphylokinase derivatives with reduced immunogenicity, used for,
PT e.g. treatment of arterial thrombosis
XX
XX
PS Claim 7; Page -: 101pp; English.
XX
XX
CC The present sequence is a specifically claimed Staphylokinase SakSTAR
CC variant. This variant has one aminoacid that has been substituted
CC by another aminoacid that reduces the reactivity with monoclonal
CC antibodies and absorption of SakSTAR-specific antibodies from plasma
CC of patients treated with staphylokinase. The derivatives can also be
CC substituted with cysteine modified with PEG to maintain the specific
CC activity and significantly reduce the plasma clearance. They have
CC altered immunogenicity without markedly reducing the specific activity.
CC This sequence has anti-thrombotic, cardiant activity and a fibrinolytic
CC potency in human plasma. The new staphylokinase derivatives are used
CC for treatment of arterial thrombosis, especially myocardial infarction.
CC The compositions can be used in human or veterinary practice.
CC Note: The present sequence is not shown in the specification, but is
CC derived from the Staphylococcus aureus wild type staphylokinase sequence
CC given in figure 1.
XX
XX
SQ Sequence 136 AA;

```

Query Match 96.3%; Score 131; DB 20; Length 136;  
Best Local Similarity 100.0%; Pred. No. 5e-126;  
Matches 131; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 SSSFDKGYKKGDDASYPEPTGPIYLAWNTGVDSKGNELSHYVEFPKPKTITLTKKI 60
    |||||||
Db 1 sssfdkgykkygddasyfepitgpylmvntgvdsksynelshpyvefpikpkitltkkl 60
OY 61 EYVVEALDATATYKRFRAVELDPSAKIEVYTDKNKKKEETSPFTEKGFAVVDLSEKI 120
    |||||||
Db 61 eyvvealdatakykefrvaeldpssakievtytdknkkkeetksfpftekgyfvvdlselh 120

```

```

OY 121 KNPGFNLITKV 131
    |||||||
Db 121 knpgfnliltkv 131

RESULT 10
AAV15026
ID AAV15026 standard; Protein: 136 AA.
XX
XX
AC AAV15026;
XX
XX
DT 03-NOV-1999 (first entry)
XX
XX
DE Staphylokinase variant SakSTAR (V132N).
XX
XX
KW Staphylokinase; variant; immunogenic; specificity; derivative; cysteine;
KW SakSTAR (V132N); anti-thrombotic; fibrinolytic; cardiant; veterinary;
KW Staphylococcus aureus; myocardial infarction; arterial thrombosis;
KW monoclonal antibody; polyethylene glycol; PEG; plasma clearance.
XX
OS Staphylococcus aureus.
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 132 /note="Wild type Val is substituted by Asn"
FT
XX
XX
PN MO9940198-A2.
XX
XX
PD 12-AUG-1999.
XX
XX
PE 04-FEB-1999; 99WO-EP00748.
XX
XX
PR 06-FEB-1998; 98EP-0200365.
PR 04-FEB-1998; 98EP-0200323.
XX
XX
PA (COLL/) COLLEN D J.
PA (LEUV-) LEUVEN RES & DEV VZW.
XX
XX
PI Collen DJ;
XX
XX
DR WPI: 1999-508504/42.
XX
XX
PT Staphylokinase derivatives with reduced immunogenicity, used for,
PT e.g. treatment of arterial thrombosis
XX
XX
PS Claim 7; Page -: 101pp; English.
XX
XX
CC The present sequence is a specifically claimed Staphylokinase SakSTAR
CC variant. This variant has one aminoacid that has been substituted
CC by another aminoacid that reduces the reactivity with monoclonal
CC antibodies and absorption of SakSTAR-specific antibodies from plasma
CC of patients treated with staphylokinase. The derivatives can also be
CC substituted with cysteine modified with PEG to maintain the specific
CC activity and significantly reduce the plasma clearance. They have
CC altered immunogenicity without markedly reducing the specific activity.
CC This sequence has anti-thrombotic, cardiant activity and a fibrinolytic
CC potency in human plasma. The new staphylokinase derivatives are used
CC for treatment of arterial thrombosis, especially myocardial infarction.
CC The compositions can be used in human or veterinary practice.
CC Note: The present sequence is not shown in the specification, but is
CC derived from the Staphylococcus aureus wild type staphylokinase sequence
CC given in figure 1.
XX
XX
SQ Sequence 136 AA;

```

Query Match 96.3%; Score 131; DB 20; Length 136;  
Best Local Similarity 100.0%; Pred. No. 5e-126;  
Matches 131; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 SSSFDKGYKKGDDASYPEPTGPIYLAWNTGVDSKGNELSHYVEFPKPKTITLTKKI 60

```

```

Db      1 sssfdkgykkggdasyfeptgpylmvntgvdskgnellsphvvefpikpgtltlckek1 60
QY      61 EYVEMALDATALYKEFRVVELDPSAKIEVYYRNKKKKKEETKSPITTEKGFVVPDLSEHI 120
Db      61 eyvewaldatalykefrvveldpsakievlydnkkkkektsfpittekgtfvvpdlsehi 120
QY      121 KNPGFNLTIKV 131
Db      121 knpgfnliltkv 131

RESULT 11
AA15027
ID      AA15027 standard; Protein; 136 AA.
AC      AA15027;
XX
XX      03-NOV-1999 (first entry)
DT
DE      Staphylokinase variant SakSTAR (V132R).
XX
XX      Staphylokinase; variant; immunogenic; specificity; derivative; cysteine;
KM      SakSTAR (V132R); anti-thrombotic; fibrinolytic; cardiant; veterinary;
KM      Staphylococcus aureus; myocardial infarction; arterial thrombosis;
KM      monoclonal antibody; polyethylene glycol; PEG; plasma clearance.
XX
XX      Staphylococcus aureus.
OS      Synthetic.
XX
XX      Key Location/Qualifiers
FH      Misc-difference 132 /note= "Wild type Val is substituted by Arg"
FT
XX
XX      WO940198-A2.
XX
XX      12-AUG-1999.
XX
XX      PF 04-FEB-1999; 99WO-EP00748.
XX
XX      PR 06-FEB-1998; 98EP-0200365.
XX      PR 04-FEB-1998; 98EP-0200323.
XX
XX      PA (COLL/) COLLEN D J.
XX      PA (LEUV-) LEUVEN RES & DEV VZW.
XX
XX      PI Collen DJ;
XX
XX      WPI; 1999-508504/42.
XX
XX      Staphylokinase derivatives with reduced immunogenicity, used for,
PT      e.g. treatment of arterial thrombosis
XX
XX      Claim 7; Page -: 101pp; English.
XX
XX      The present sequence is a specifically claimed Staphylokinase SakSTAR
CC      variant. This variant has one aminoacid that has been substituted
CC      by another aminoacid that reduces the reactivity with monoclonal
CC      antibodies and absorption of SakSTAR-specific antibodies from plasma
CC      of patients treated with staphylokinase. The derivatives can also be
CC      substituted with cysteine modified with PEG to maintain the specific
CC      activity and significantly reduce the plasma clearance. They have
CC      altered immunogenicity without markedly reducing the specific activity.
CC      This sequence has anti-thrombotic, cardiant activity and a fibrinolytic
CC      potency in human plasma. The new staphylokinase derivatives are used
CC      for treatment of arterial thrombosis, especially myocardial infarction.
CC      The compositions can be used in human or veterinary practice.
CC      Note: The present sequence is not shown in the specification, but is
CC      derived from the Staphylococcus aureus wild type staphylokinase sequence
CC      given in Figure 1.
XX
XX      Sequence 136 AA;
SQ
```

```

Query Match          96.3%; Score 131; DB 20; Length 136;
Best Local Similarity 100.0%; Pred. No. 5e-126;
Matches 131; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSSFDKGYKKGGDASYFEPTGPYLMVNTGVDKGNELLSPHYVEFPKPGTTLTKEK1 60
Db      1 sssfdkgykkggdasyfeptgpylmvntgvdskgnellsphvvefpikpgtltlckek1 60
QY      61 EYVEMALDATALYKEFRVVELDPSAKIEVYYRNKKKKKEETKSPITTEKGFVVPDLSEHI 120
Db      61 eyvewaldatalykefrvveldpsakievlydnkkkkektsfpittekgtfvvpdlsehi 120
QY      121 KNPGFNLTIKV 131
Db      121 knpgfnliltkv 131

RESULT 12
AA15016
ID      AA15016 standard; Protein; 136 AA.
AC      AA15016;
XX
XX      03-NOV-1999 (first entry)
DT
DE      Staphylokinase variant SakSTAR (V132A).
XX
XX      Staphylokinase; variant; immunogenic; specificity; derivative; cysteine;
KM      SakSTAR (V132A); anti-thrombotic; fibrinolytic; cardiant; veterinary;
KM      Staphylococcus aureus; myocardial infarction; arterial thrombosis;
KM      monoclonal antibody; polyethylene glycol; PEG; plasma clearance.
XX
XX      Staphylococcus aureus.
OS      Synthetic.
XX
XX      Key Location/Qualifiers
FH      Misc-difference 132 /note= "Wild type Val is substituted by Ala"
FT
XX
XX      WO940198-A2.
XX
XX      12-AUG-1999.
XX
XX      PF 04-FEB-1999; 99WO-EP00748.
XX
XX      PR 06-FEB-1998; 98EP-0200365.
XX      PR 04-FEB-1998; 98EP-0200323.
XX
XX      PA (COLL/) COLLEN D J.
XX      PA (LEUV-) LEUVEN RES & DEV VZW.
XX
XX      PI Collen DJ;
XX
XX      WPI; 1999-508504/42.
XX
XX      Staphylokinase derivatives with reduced immunogenicity, used for,
PT      e.g. treatment of arterial thrombosis
XX
XX      Claim 7; Page -: 101pp; English.
XX
XX      The present sequence is a specifically claimed Staphylokinase SakSTAR
CC      variant. This variant has one aminoacid that has been substituted
CC      by another aminoacid that reduces the reactivity with monoclonal
CC      antibodies and absorption of SakSTAR-specific antibodies from plasma
CC      of patients treated with staphylokinase. The derivatives can also be
CC      substituted with cysteine modified with PEG to maintain the specific
CC      activity and significantly reduce the plasma clearance. They have
CC      altered immunogenicity without markedly reducing the specific activity.
CC      This sequence has anti-thrombotic, cardiant activity and a fibrinolytic
CC      potency in human plasma. The new staphylokinase derivatives are used
CC      for treatment of arterial thrombosis, especially myocardial infarction.
CC      The compositions can be used in human or veterinary practice.
CC
```



CC Note: The present sequence is not shown in the specification, but is  
CC derived from the Staphylococcus aureus wild type staphylokinase sequence  
CC given in figure 1.

CC Sequence 136 AA;

Query Match 96.3%; Score 131; DB 20; Length 136;  
Best Local Similarity 100.0%; Pred. No. 5e-126;  
Matches 131; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSSPFGKYYKKGDDASYPETGPIYLVWVTGVDSKGNELSPHYVEPIKPGTTLTKRKI 60  
DB 1 sssstkgkykkgddasyfepcgyilmvntgvdskgnelisphvefpikpgtltkexi 60  
OY 61 EYVEMADATAYKEFRVVELDPSAKIEVTYYDKNKKKEETKSPITKGFVVPDLSEHI 120  
DB 61 eyvewaldataykefrvveldpesakievtydknkkkeetksfptekgfvvpdlsehi 120  
OY 121 KNPGFNLTIKV 131  
DB 121 knpgfnliltkv 131

RESULT 13

AAW03087  
ID AAW03087 standard; protein; 136 AA.

XX AC AAW03087;

XX DT 19-FEB-1997 (first entry)

XX DE Staphylokinase derivative M20 with altered charge cluster 20.

KW Staphylokinase; mutant; mutein; variant; immunogenicity: decrease;  
KW derivative; SakSTAR; arterial thrombosis; thrombolytic agent.

XX OS Staphylococcus aureus.

OS Synthetic.

XX FH Location/Qualifiers

FT Key 5.6

FT Region /label= altered\_M20\_cluster  
/note= "wild-type cluster 20 sequence Asp-Lys  
has been replaced by Ala-Ala to  
reduce immunogenicity"

FT FT reduce immunogenicity"

XX PN W09621016-A2.

XX PD 11-JUL-1996.

XX PF 03-JAN-1996; 96WO-EP00081.

XX PR 17-NOV-1995; 95JP-0299781.

XX PR 06-JAN-1995; 95EP-0200023.

XX PR 11-JAN-1995; 95US-0371505.

XX PR 09-JUN-1995; 95EP-0201531.

XX PR 06-JUL-1995; 95US-0499092.

XX PA (COLL/) COLLEN D.

XX PA (LEUV-) LEUVEN RES & DEV VZW.

XX PI Colleen D;

XX DR WPI; 1996-333991/33.

XX PT New staphylokinase derivs. having reduced immunogenicity - useful  
XX for treating arterial thrombosis  
XX Example 2; Page -: 58pp; English.  
XX Staphylokinase derivatives showing a reduced immunogenicity as  
CC compared to wild-type staphylokinase are claimed. The derivatives

CC are useful as thrombolytic agents to treat arterial thrombosis and  
CC are pref. produced by eliminating at least one of the epitopes  
CC indicated in the features table. The epitopes are destroyed by  
CC replacing one or more amino acid residues in a charge cluster by an  
CC Ala residue. Mutations are introduced using site-directed  
CC mutagenesis on wild-type staphylokinase genomic DNA from the S. aureus  
CC lysogenic strain SakSTAR. The present sequence is a specific  
CC example of a staphylokinase derivative according to the invention;  
CC the sequence does not appear in the specification but can be generated  
CC from the wild-type staphylokinase sequence using the description given  
CC in Table 1 of the disclosure.

XX Sequence 136 AA;

Query Match 95.6%; Score 130; DB 17; Length 136;  
Best Local Similarity 100.0%; Pred. No. 5.2e-125;  
Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GKYYKGGDASYPETGPIYLVWVTGVDSKGNELSPHYVEPIKPGTTLTKRKIYYEW 66  
DB 7 gkykkgddasyfepcgyilmvntgvdskgnelisphvefpikpgtltkexiyyew 66

OY 67 ALDATAYKEFRVVELDPSAKIEVTYYDKNKKKEETKSPITKGFVVPDLSEHIKNPGFN 126  
DB 67 aldataykefrvveldpesakievtydknkkkeetksfptekgfvvpdlsehi knpgfn 126

OY 127 LITKVVIEKK 136  
DB 127 litkvviekk 136

RESULT 14

AAW03101  
ID AAW03101 standard; protein; 136 AA.

XX AC AAW03101;

XX DT 19-FEB-1997 (first entry)

XX DE Staphylokinase derivative M18 with altered charge cluster 18.

KW Staphylokinase; mutant; mutein; variant; immunogenicity: decrease;  
KW derivative; SakSTAR; arterial thrombosis; thrombolytic agent.

XX OS Staphylococcus aureus.

OS Synthetic.

XX FH Location/Qualifiers

FT Key 130

FT Region /label= altered\_M18\_cluster  
/note= "wild-type cluster 18 sequence Lys  
has been replaced by Ala to reduce  
immunogenicity"

FT FT immunogenicity"

XX PN W09621016-A2.

XX PD 11-JUL-1996.

XX PF 03-JAN-1996; 96WO-EP00081.

XX PR 17-NOV-1995; 95JP-0299781.

XX PR 06-JAN-1995; 95EP-0200023.

XX PR 11-JAN-1995; 95US-0371505.

XX PR 09-JUN-1995; 95EP-0201531.

XX PR 06-JUL-1995; 95US-0499092.

XX PA (COLL/) COLLEN D.  
XX PA (LEUV-) LEUVEN RES & DEV VZW.  
XX PI Colleen D;  
XX DR WPI; 1996-333991/33.

Query Match	Best local Similarity	Score 129; DB 17; Length 136;
Match 129; Conservative	100.0%; Pred. No. 5.5e-124;	Mismatches 0; Indels 0; Gaps 0.
QY 1 SSSFPKGRYKKRDKDDASYEPTGPYIMVNVNTGDSKGNELLSPHYVEFPKPGTTLTKRKI 60		
Db 1 sssfckkygkkgkddasyteptgpylmvntvgdskgnellsphyvefpkpgtltlkrki 60		
QY 61 EYVEMALDATAKKEFRRVELDPSAKIEVTTYDKNNKKEETKSPITEKGFVVPDLSEHI 120		
Db 61 eyvewaldataykefrvelpsakievttydknnkkeetksfpitekgfvvpdlsehl 120		
QY 121 KNPGENLT 129		
Db 121 knpgfnlilt 129		
RESULT 15		
AAV15023		
ID AAV15023 standard; Protein; 136 AA.		
AC AAV15023;		
DT 03-NOV-1999 (first entry)		
DE Staphylokinase variant SakSTAR (K130T).		
XX Staphylokinase; variant; immunogenic; specificity; derivative; cysteine; SakSTAR (K130T); anti-thrombotic; fibrinolytic; cardiant; veterinary; KW Staphylococcus aureus; myocardial infarction; arterial thrombosis; monoclonal antibody; polyethylene glycol; PEG; plasma clearance.		
XX Staphylococcus aureus.		
OS Synthetic.		
XX Key	Location/Qualifiers	
FT Misc-difference 130		
PN W09940198-A2.	/note= "Wild type Lys is substituted by Thr"	
XX 12-AUG-1999.		
XX 04-FEB-1999;	99WO-EP00748.	
XX 06-FEB-1998;	98EP-0200365.	
PR 04-FEB-1998;	98EP-0200323.	
XX (COLL/) COLLEN D J.		
XX (LEUV-) LEUVEN RES & DEV VZW.		

xx		Collen DJ;	
pi			
xx	wpt; 1999-508504/42.		
xx			
xx		Staphylokinase derivatives with reduced immunogenicity, used for,	
pr	e.g. treatment of arterial thrombosis		
xx			
ps	Claim 7; Page -; 101pp; English.		
cc	The present sequence is a specifically claimed Staphylokinase SakSTAR		
cc	variant. This variant has one aminoacid that has been substituted		
cc	by another aminoacid that reduces the reactivity with monoclonal		
cc	antibodies and absorption of SakSTAR-specific antibodies from plasma		
cc	of patients treated with staphylokinase. The derivatives can also be		
cc	substituted with cysteine modified with PEG to maintain the specific		
cc	activity and significantly reduce the plasma clearance. They have		
cc	altered immunogenicity without markedly reducing the specific activity.		
cc	This sequence has anti-thrombotic, cardiant activily and a fibrinolytic		
cc	potency in human plasma. The new staphylokinase derivatives are used		
cc	for treatment of arterial thrombosis, especially myocardial infarction.		
cc	The compositions can be used in human or veterinary practice.		
cc	Note: The present sequence is not shown in the specification, but is		
cc	derived from the Staphylococcus aureus wild type staphylokinase sequence		
cc	given in figure 1.		
xx			
xx			
sQ	Sequence      136 AA;		
	Query Match                 94.9%; Score 129; DB 20; Length 136; Best Local Similarity 100.0%; Pred. No. 5.5e-124; Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
OY	1 SSSFFKGRKKRDKDASYPEPTGPRILAMVNTGVDSKNELLSPHYVEFPIRKGTTLTKETI 60       		
Dd	1 sssIdkgykkxgdassyletpgpylmvntgvdsksynellsphyvefipikpgtltlIketi 60       		
OY	61 EYYEWALDATAYAKFEFRVELDPISAKIEITYDYDNKKKEETSFPTEKGFWPDLSEHI 120       		
Dd	61 eyyewaldatayakfefrveldpisakietydydnkkkeetsfpItetkgfwvpdlsehl 120       		
OY	121 KNPGNLTIT 129       		
Dd	121 knpgfnilt 129       		

Search completed: April 22, 2002, 10:45:23  
Job time: 208 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 10:43:45 ; Search time 12.66 Seconds

(without alignments)  
241,742 Million cell updates/sec

Title: US-09-601-490-1

Perfect score: 136

Sequence: 1 SSSFDKXGKXKGGDASYPEP.....SEHIKNPGLITKVIIEK 136

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 212252 seqs, 22503292 residues

Word size : 0

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PCITUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	136	100.0	136	1 US-08-371-505-2	Sequence 2, Appli
2	136	100.0	136	2 US-08-784-971-5	Sequence 5, Appli
3	136	100.0	137	1 US-08-256-261-6	Sequence 6, Appli
4	136	100.0	137	3 US-08-852-289-6	Sequence 6, Appli
5	102	75.0	137	1 US-08-256-261-4	Sequence 4, Appli
6	102	75.0	137	3 US-08-852-299-4	Sequence 4, Appli
7	93	68.4	123	1 US-08-256-261-10	Sequence 10, Appli
8	93	68.4	123	3 US-08-852-299-10	Sequence 10, Appli
9	93	68.4	127	1 US-08-256-261-8	Sequence 8, Appli
10	93	68.4	127	3 US-08-852-299-8	Sequence 8, Appli
11	93	68.4	137	1 US-08-256-261-2	Sequence 2, Appli
12	93	68.4	137	1 US-08-256-261-12	Sequence 12, Appli
13	93	68.4	137	1 US-08-256-261-14	Sequence 14, Appli
14	93	68.4	137	3 US-08-852-299-2	Sequence 2, Appli
15	93	68.4	137	3 US-08-852-289-12	Sequence 12, Appli
16	93	68.4	137	3 US-08-852-299-14	Sequence 14, Appli
17	93	68.4	163	1 US-08-256-261-17	Sequence 17, Appli
18	93	68.4	163	3 US-08-852-299-17	Sequence 17, Appli
19	82	60.3	126	1 US-08-075-545-2	Sequence 2, Appli
20	82	60.3	163	1 US-08-075-545-1	Sequence 1, Appli
21	82	60.3	163	1 US-08-934-222-21	Sequence 21, Appli
22	82	60.3	163	1 US-08-934-222-22	Sequence 22, Appli
23	82	60.3	163	1 US-08-934-222-23	Sequence 23, Appli
24	82	60.3	163	1 US-08-934-222-24	Sequence 24, Appli
25	82	60.3	163	1 US-08-934-222-25	Sequence 25, Appli
26	82	60.3	163	1 US-08-934-222-26	Sequence 26, Appli
27	82	60.3	163	1 US-08-934-222-27	Sequence 27, Appli

28	12	8.8	12	2 US-08-532-818-22	Sequence 22, Appli
29	12	8.8	12	3 US-09-231-797-21	Sequence 21, Appli
30	12	8.8	12	3 US-09-231-797-22	Sequence 22, Appli
31	12	8.8	12	3 US-08-934-224-21	Sequence 21, Appli
32	12	8.8	12	3 US-08-934-224-22	Sequence 22, Appli
33	12	8.8	12	3 US-08-934-224-23	Sequence 23, Appli
34	12	8.8	12	3 US-08-934-224-24	Sequence 24, Appli
35	12	8.8	12	3 US-08-934-224-25	Sequence 25, Appli
36	12	8.8	12	3 US-08-934-224-26	Sequence 26, Appli
37	12	8.8	12	3 US-08-934-224-27	Sequence 27, Appli
38	12	8.8	12	3 US-08-934-224-28	Sequence 28, Appli
39	12	8.8	12	3 US-08-934-224-29	Sequence 29, Appli
40	12	8.8	12	3 US-08-934-224-30	Sequence 30, Appli
41	12	8.8	12	3 US-08-934-224-31	Sequence 31, Appli
42	12	8.8	12	3 US-08-934-224-32	Sequence 32, Appli
43	12	8.8	12	3 US-08-934-224-33	Sequence 33, Appli
44	12	8.8	12	3 US-08-934-224-34	Sequence 34, Appli
45	12	8.8	12	3 US-08-934-224-35	Sequence 35, Appli
46	12	8.8	12	3 US-08-934-224-36	Sequence 36, Appli
47	12	8.8	12	3 US-08-934-224-37	Sequence 37, Appli
48	12	8.8	12	3 US-08-934-224-38	Sequence 38, Appli
49	12	8.8	12	3 US-08-934-224-39	Sequence 39, Appli
50	12	8.8	12	3 US-08-934-224-40	Sequence 40, Appli
51	12	8.8	12	3 US-08-934-224-41	Sequence 41, Appli
52	12	8.8	12	3 US-08-934-224-42	Sequence 42, Appli
53	12	8.8	12	3 US-08-934-224-43	Sequence 43, Appli
54	12	8.8	12	3 US-08-934-224-44	Sequence 44, Appli
55	12	8.8	12	3 US-08-934-224-45	Sequence 45, Appli
56	12	8.8	12	3 US-08-934-224-46	Sequence 46, Appli
57	12	8.8	12	3 US-08-934-224-47	Sequence 47, Appli
58	12	8.8	12	3 US-08-934-224-48	Sequence 48, Appli
59	12	8.8	12	3 US-08-934-224-49	Sequence 49, Appli
60	12	8.8	12	3 US-08-934-224-50	Sequence 50, Appli
61	12	8.8	12	3 US-08-934-224-51	Sequence 51, Appli
62	12	8.8	12	3 US-08-934-224-52	Sequence 52, Appli
63	12	8.8	12	3 US-08-934-224-53	Sequence 53, Appli
64	12	8.8	12	3 US-08-934-224-54	Sequence 54, Appli
65	12	8.8	12	3 US-08-934-224-55	Sequence 55, Appli
66	12	8.8	12	3 US-08-934-224-56	Sequence 56, Appli
67	12	8.8	12	3 US-08-934-224-57	Sequence 57, Appli
68	12	8.8	12	3 US-08-934-224-58	Sequence 58, Appli
69	12	8.8	12	3 US-08-934-224-59	Sequence 59, Appli
70	12	8.8	12	3 US-08-934-224-60	Sequence 60, Appli
71	12	8.8	12	3 US-08-934-224-61	Sequence 61, Appli
72	12	8.8	12	3 US-08-934-224-62	Sequence 62, Appli
73	12	8.8	12	3 US-08-934-224-63	Sequence 63, Appli
74	12	8.8	12	3 US-08-934-224-64	Sequence 64, Appli
75	12	8.8	12	3 US-08-934-224-65	Sequence 65, Appli
76	12	8.8	12	3 US-08-934-224-66	Sequence 66, Appli
77	12	8.8	12	3 US-08-934-224-67	Sequence 67, Appli
78	12	8.8	12	3 US-08-934-224-68	Sequence 68, Appli
79	12	8.8	12	3 US-08-934-224-69	Sequence 69, Appli
80	12	8.8	12	3 US-08-934-224-70	Sequence 70, Appli
81	12	8.8	12	3 US-08-934-224-71	Sequence 71, Appli
82	12	8.8	12	3 US-08-934-224-72	Sequence 72, Appli
83	12	8.8	12	3 US-08-934-224-73	Sequence 73, Appli
84	12	8.8	12	3 US-08-934-224-74	Sequence 74, Appli
85	12	8.8	12	3 US-08-934-224-75	Sequence 75, Appli
86	12	8.8	12	3 US-08-934-224-76	Sequence 76, Appli
87	12	8.8	12	3 US-08-934-224-77	Sequence 77, Appli
88	12	8.8	12	3 US-08-934-224-78	Sequence 78, Appli
89	12	8.8	12	3 US-08-934-224-79	Sequence 79, Appli
90	12	8.8	12	3 US-08-934-224-80	Sequence 80, Appli
91	12	8.8	12	3 US-08-934-224-81	Sequence 81, Appli
92	12	8.8	12	3 US-08-934-224-82	Sequence 82, Appli
93	12	8.8	12	3 US-08-934-224-83	Sequence 83, Appli
94	12	8.8	12	3 US-08-934-224-84	Sequence 84, Appli
95	12	8.8	12	3 US-08-934-224-85	Sequence 85, Appli
96	12	8.8	12	3 US-08-934-224-86	Sequence 86, Appli
97	12	8.8	12	3 US-08-934-224-87	Sequence 87, Appli
98	12	8.8	12	3 US-08-934-224-88	Sequence 88, Appli
99	12	8.8	12	3 US-08-934-224-89	Sequence 89, Appli
100	12	8.8	12	3 US-08-934-224-90	Sequence 90, Appli



247	5	3.7	19	1	US-08-165-038-6	Sequence 6, Appli	320	5	3.7	29	5	PCT-US95-04589-17	Sequence 17, Appli
248	5	3.7	19	2	US-08-876-781-6	Sequence 6, Appli	321	5	3.7	33	4	US-08-836-252A-1	Sequence 1, Appli
249	5	3.7	21	1	US-07-807-529A-17	Sequence 17, Appli	322	5	3.7	33	4	US-09-082-279B-1105	Sequence 1105, Ap
250	5	3.7	21	3	US-08-300-928C-66	Sequence 66, Appli	323	5	3.7	33	4	US-09-082-279B-1106	Sequence 1106, Ap
251	5	3.7	21	3	US-08-430-944D-66	Sequence 66, Appli	324	5	3.7	34	4	US-09-082-279B-1107	Sequence 1107, Ap
252	5	3.7	21	3	US-08-430-014-66	Sequence 66, Appli	325	5	3.7	35	4	US-09-082-279B-1100	Sequence 1100, Ap
253	5	3.7	22	1	US-08-431-184-66	Sequence 66, Appli	326	5	3.7	35	4	US-09-082-279B-1102	Sequence 1102, Ap
254	5	3.7	22	1	US-08-343-943-7	Sequence 7, Appli	327	5	3.7	35	4	US-09-082-279B-1115	Sequence 1115, Ap
255	5	3.7	22	3	US-08-940-095-9	Sequence 9, Appli	328	5	3.7	35	4	US-09-082-279B-1158	Sequence 1158, Ap
256	5	3.7	22	3	US-08-940-093-9	Sequence 9, Appli	329	5	3.7	35	4	US-09-082-279B-1116	Sequence 1116, Ap
257	5	3.7	22	3	US-08-940-096-9	Sequence 9, Appli	330	5	3.7	35	4	US-09-082-279B-1185	Sequence 1185, Ap
258	5	3.7	22	4	US-09-082-279B-1292	Sequence 1292, Ap	331	5	3.7	35	4	US-09-082-279B-1186	Sequence 1186, Ap
259	5	3.7	22	4	US-08-082-279B-1327	Sequence 1327, Ap	332	5	3.7	35	4	US-09-082-279B-1311	Sequence 1311, Ap
260	5	3.7	22	4	US-09-465-719-9	Sequence 9, Appli	333	5	3.7	35	6	5190919-5	Patent No. 5190919
261	5	3.7	23	2	US-08-637-759B-242	Sequence 242, App	334	5	3.7	35	6	5190919-6	Sequence 1072, Ap
262	5	3.7	23	3	US-08-871-355A-242	Sequence 242, App	335	5	3.7	36	4	US-09-082-279B-1072	Sequence 1073, Ap
263	5	3.7	23	4	US-09-082-279B-1206	Sequence 1206, Ap	336	5	3.7	36	4	US-09-082-279B-1073	Patent No. 5190919
264	5	3.7	23	4	US-09-082-279B-1207	Sequence 1207, Ap	337	5	3.7	36	6	5190919-15	Patent No. 5190919
265	5	3.7	23	4	US-09-082-279B-1288	Sequence 1289, Ap	338	5	3.7	37	4	US-08-836-252A-4	Sequence 4, Appli
266	5	3.7	23	4	US-09-082-279B-1290	Sequence 1290, Ap	339	5	3.7	37	4	US-09-082-279B-1117	Sequence 1117, Ap
267	5	3.7	23	4	US-09-082-279B-1291	Sequence 1291, Ap	340	5	3.7	39	4	US-09-082-279B-1116	Sequence 1116, Ap
268	5	3.7	23	4	US-09-082-279B-1293	Sequence 1293, Ap	341	5	3.7	39	4	US-09-082-279B-1130	Sequence 1130, Ap
269	5	3.7	23	4	US-09-082-279B-1294	Sequence 1294, Ap	342	5	3.7	39	4	US-09-082-279B-1131	Sequence 1131, Ap
270	5	3.7	23	4	US-09-082-279B-1324	Sequence 1324, Ap	343	5	3.7	39	4	US-09-082-279B-1134	Sequence 1134, Ap
271	5	3.7	23	4	US-09-082-279B-1325	Sequence 1325, Ap	344	5	3.7	39	4	US-09-082-279B-1149	Sequence 1149, Ap
272	5	3.7	23	4	US-09-082-279B-1326	Sequence 1326, Ap	345	5	3.7	39	4	US-09-082-279B-1150	Sequence 1150, Ap
273	5	3.7	23	4	US-09-082-279B-1328	Sequence 1328, Ap	346	5	3.7	40	4	US-09-082-279B-1133	Sequence 1133, Ap
274	5	3.7	23	4	US-09-082-279B-1329	Sequence 1329, Ap	347	5	3.7	40	5	PCT-US93-05640-12	Sequence 12, Appli
275	5	3.7	24	6	5190919-16	Patent No. 5190919	348	5	3.7	40	5	PCT-US93-05640-13	Sequence 23, Appli
276	5	3.7	26	1	US-07-807-529A-9	Sequence 9, Appli	349	5	3.7	41	4	US-09-082-279B-1135	Sequence 1135, Ap
277	5	3.7	26	1	US-07-942-245-231	Sequence 231, App	350	5	3.7	42	4	US-09-082-279B-1132	Sequence 1132, Ap
278	5	3.7	26	1	US-07-942-245-232	Sequence 232, App	351	5	3.7	43	1	US-08-178-477B-22	Sequence 22, Appli
279	5	3.7	26	1	US-07-942-245-233	Sequence 233, App	352	5	3.7	43	2	US-08-210-762E-6	Sequence 6, Appli
280	5	3.7	26	1	US-07-942-245-234	Sequence 234, App	353	5	3.7	43	2	US-08-210-762E-17	Sequence 17, Appli
281	5	3.7	26	1	US-07-942-245-235	Sequence 235, App	354	5	3.7	43	2	US-08-488-161-38	Sequence 38, Appli
282	5	3.7	26	1	US-07-942-245-237	Sequence 237, App	355	5	3.7	43	3	US-09-273-685-38	Sequence 38, Appli
283	5	3.7	26	1	US-07-942-245-238	Sequence 238, App	356	5	3.7	43	4	US-09-015-030-7	Sequence 7, Appli
284	5	3.7	26	1	US-07-942-245-239	Sequence 239, App	357	5	3.7	43	5	PCT-US95-11334-38	Sequence 38, Appli
285	5	3.7	26	1	US-07-942-245-240	Sequence 240, App	358	5	3.7	44	4	US-09-046-894-41	Sequence 41, Appli
286	5	3.7	26	1	US-07-942-245-241	Sequence 241, App	359	5	3.7	44	4	US-09-046-894-42	Sequence 42, Appli
287	5	3.7	26	1	US-07-942-245-244	Sequence 244, App	360	5	3.7	45	1	US-08-664-856A-2	Sequence 2, Appli
288	5	3.7	26	1	US-07-942-245-247	Sequence 247, App	361	5	3.7	45	1	US-08-861-877-2	Sequence 2, Appli
289	5	3.7	26	1	US-07-942-245-248	Sequence 248, App	362	5	3.7	45	2	US-08-931-877-2	Sequence 2, Appli
290	5	3.7	26	1	US-07-942-245-249	Sequence 249, App	363	5	3.7	45	2	US-08-664-857A-2	Sequence 2, Appli
291	5	3.7	26	1	US-07-942-245-250	Sequence 250, App	364	5	3.7	45	3	US-09-069-484-2	Sequence 2, Appli
292	5	3.7	26	1	US-07-942-245-252	Sequence 252, App	365	5	3.7	45	4	US-09-135-599-2	Sequence 2, Appli
293	5	3.7	26	1	US-07-942-245-253	Sequence 253, App	366	5	3.7	45	4	US-08-905-223-367	Sequence 367, App
294	5	3.7	26	1	US-07-942-245-254	Sequence 254, App	367	5	3.7	45	4	US-08-711-417C-179	Sequence 179, App
295	5	3.7	26	1	US-07-942-245-260	Sequence 260, App	368	5	3.7	45	4	US-08-711-417C-181	Sequence 181, App
296	5	3.7	26	1	US-07-942-245-261	Sequence 261, App	369	5	3.7	45	4	US-09-369-744-2	Sequence 2, Appli
297	5	3.7	26	1	US-07-942-245-269	Sequence 269, App	370	5	3.7	50	1	US-07-956-700B-55	Sequence 55, Appli
298	5	3.7	26	1	US-07-942-245-424	Sequence 424, App	371	5	3.7	50	1	US-08-476-537-55	Sequence 55, Appli
299	5	3.7	26	1	US-07-942-245-426	Sequence 426, App	372	5	3.7	50	1	US-08-485-607-55	Sequence 55, Appli
300	5	3.7	26	1	US-07-942-245-427	Sequence 427, App	373	5	3.7	50	2	US-08-475-879-55	Sequence 55, Appli
301	5	3.7	26	1	US-07-942-245-429	Sequence 429, App	374	5	3.7	54	1	US-08-165-038-1	Sequence 1, Appli
302	5	3.7	26	1	US-07-942-245-433	Sequence 433, App	375	5	3.7	54	1	US-08-621-081A-1	Sequence 1, Appli
303	5	3.7	26	1	US-07-942-245-434	Sequence 434, App	376	5	3.7	54	2	US-08-876-781-1	Sequence 1, Appli
304	5	3.7	26	1	US-07-942-245-437	Sequence 437, App	377	5	3.7	56	5	PCT-US92-08558-6	Sequence 6, Appli
305	5	3.7	26	1	US-07-942-245-445	Sequence 445, App	378	5	3.7	59	4	US-09-117-715-11	Sequence 11, Appli
306	5	3.7	26	1	US-07-942-245-446	Sequence 446, App	379	5	3.7	69	4	US-09-117-715-11	Sequence 11, Appli
307	5	3.7	26	1	US-07-942-245-447	Sequence 447, App	380	5	3.7	70	2	US-08-025-151-16	Sequence 16, App
308	5	3.7	26	1	US-07-942-245-449	Sequence 449, App	381	5	3.7	72	4	US-08-210-762E-16	Sequence 16, Appli
309	5	3.7	26	1	US-07-942-245-451	Sequence 451, App	382	5	3.7	72	4	US-08-836-500A-6	Sequence 6, Appli
310	5	3.7	26	3	US-08-300-928C-19	Sequence 19, Appli	383	5	3.7	75	3	US-09-100-804-21	Sequence 21, Appli
311	5	3.7	26	3	US-08-430-944D-19	Sequence 19, Appli	384	5	3.7	78	2	US-08-469-412A-11	Sequence 11, Appli
312	5	3.7	26	3	US-08-430-014-19	Sequence 19, Appli	385	5	3.7	79	2	US-09-021-715-11	Sequence 11, Appli
313	5	3.7	26	3	US-08-431-184-19	Sequence 19, Appli	386	5	3.7	79	4	US-08-469-412A-12	Sequence 12, Appli
314	5	3.7	26	5	PCT-US93-02462-9	Sequence 9, Appli	387	5	3.7	79	4	US-09-021-715-12	Sequence 12, Appli
315	5	3.7	28	4	US-09-248-588-41	Sequence 41, Appli	388	5	3.7	79	4	US-08-225-575B-8	Sequence 8, Appli
316	5	3.7	29	2	US-08-310-912A-17	Sequence 17, Appli	389	5	3.7	81	1	US-09-100-802-5	Sequence 5, Appli
317	5	3.7	29	3	US-08-841-089-17	Sequence 17, Appli	390	5	3.7	82	3	US-08-300-928C-16	Sequence 16, Appli
318	5	3.7	29	4	US-09-301-085-17	Sequence 17, Appli	391	5	3.7	82	3	US-08-430-944D-16	Sequence 16, Appli
319	5	3.7	29	5	PCT-US95-04570-17	Sequence 17, Appli	392	5	3.7	82	3	US-08-430-014-16	Sequence 16, Appli

393	5	3.7	82	3	US-08-431-184-16	Sequence 16, Appl	466	5	3.7	111	3	US-08-430-944D-6	Sequence 6, Appl
394	5	3.7	83	1	US-07-662-193-7	Sequence 7, Appl	467	5	3.7	111	3	US-08-430-014-6	Sequence 6, Appl
395	5	3.7	83	1	US-07-662-193-8	Sequence 8, Appl	468	5	3.7	111	3	US-08-881-037-20	Sequence 20, Appl
396	5	3.7	86	2	US-08-459-568-82	Sequence 82, Appl	469	5	3.7	111	3	US-08-431-184-6	Sequence 6, Appl
397	5	3.7	86	2	US-08-399-411-82	Sequence 82, Appl	470	5	3.7	111	3	US-08-637-647-15	Sequence 15, Appl
398	5	3.7	86	3	US-08-516-859A-82	Sequence 82, Appl	471	5	3.7	111	5	PCT-US95-17025-15	Sequence 15, Appl
399	5	3.7	87	3	US-08-554-840-17	Sequence 17, Appl	472	5	3.7	112	2	US-08-672-345C-107	Sequence 107, App
400	5	3.7	87	3	US-08-554-840-18	Sequence 18, Appl	473	5	3.7	113	4	US-09-214-095D-118	Sequence 118, App
401	5	3.7	87	3	US-08-554-840-20	Sequence 20, Appl	474	5	3.7	114	3	US-08-483-749A-10	Sequence 10, Appl
402	5	3.7	90	3	US-08-300-928C-14	Sequence 14, Appl	475	5	3.7	115	3	US-08-838-682-8	Sequence 8, Appl
403	5	3.7	90	3	US-08-430-944D-14	Sequence 14, Appl	476	5	3.7	115	4	US-08-895-914-8	Sequence 8, Appl
404	5	3.7	90	3	US-08-430-014-14	Sequence 14, Appl	477	5	3.7	115	4	US-09-357-710A-8	Sequence 8, Appl
405	5	3.7	90	3	US-08-431-184-14	Sequence 14, Appl	478	5	3.7	116	4	US-07-634-278-3	Sequence 3, Appl
406	5	3.7	90	4	US-09-334-601-23	Sequence 23, Appl	479	5	3.7	116	1	US-07-634-278-14	Sequence 14, Appl
407	5	3.7	90	4	US-09-334-601-24	Sequence 24, Appl	480	5	3.7	116	1	US-08-477-728-3	Sequence 3, Appl
408	5	3.7	91	4	US-09-334-601-25	Sequence 25, Appl	481	5	3.7	116	1	US-08-477-728-14	Sequence 14, Appl
409	5	3.7	91	4	US-09-234-553-1	Sequence 1, Appl	482	5	3.7	116	1	US-08-474-040-3	Sequence 3, Appl
410	5	3.7	92	1	US-08-307-499-31	Sequence 31, Appl	483	5	3.7	116	1	US-08-474-040-14	Sequence 14, Appl
411	5	3.7	92	1	US-08-290-448A-68	Sequence 68, Appl	484	5	3.7	116	1	US-08-487-200-3	Sequence 3, Appl
412	5	3.7	92	1	US-08-290-448A-68	Sequence 68, Appl	485	5	3.7	116	1	US-08-487-200-14	Sequence 14, Appl
413	5	3.7	92	1	US-08-175-069A-68	Sequence 68, Appl	486	5	3.7	116	1	US-08-488-113B-167	Sequence 167, App
414	5	3.7	92	2	US-08-721-488-6	Sequence 6, Appl	487	5	3.7	116	1	US-08-488-113B-168	Sequence 168, App
415	5	3.7	92	3	US-08-300-928C-13	Sequence 13, Appl	488	5	3.7	116	1	US-08-477-484B-167	Sequence 167, App
416	5	3.7	92	3	US-08-430-944D-13	Sequence 13, Appl	489	5	3.7	116	1	US-08-477-484B-168	Sequence 168, App
417	5	3.7	92	3	US-08-430-014-13	Sequence 13, Appl	490	5	3.7	116	1	US-08-107-668D-53	Sequence 53, Appl
418	5	3.7	92	3	US-08-431-184-13	Sequence 13, Appl	491	5	3.7	116	1	US-08-107-668D-54	Sequence 54, Appl
419	5	3.7	92	4	US-09-299-268-31	Sequence 31, Appl	492	5	3.7	116	1	US-08-472-788A-54	Sequence 54, Appl
420	5	3.7	93	3	US-09-208-804-3	Sequence 3, Appl	493	5	3.7	116	1	US-08-472-788A-85	Sequence 85, Appl
421	5	3.7	93	3	US-08-801-743-3	Sequence 3, Appl	494	5	3.7	116	1	US-08-690-102A-4	Sequence 4, Appl
422	5	3.7	94	1	US-08-290-448A-69	Sequence 69, Appl	495	5	3.7	116	2	US-08-477-531B-53	Sequence 53, Appl
423	5	3.7	94	1	US-08-290-448A-69	Sequence 69, Appl	496	5	3.7	116	2	US-08-477-531B-54	Sequence 54, Appl
424	5	3.7	94	1	US-08-175-069A-69	Sequence 69, Appl	497	5	3.7	116	2	US-08-646-360-167	Sequence 167, App
425	5	3.7	96	1	US-07-807-529A-39	Sequence 39, Appl	498	5	3.7	116	2	US-08-646-360-168	Sequence 168, App
426	5	3.7	96	3	US-08-430-944D-103	Sequence 103, App	499	5	3.7	116	2	US-08-646-360-168	Sequence 168, App
427	5	3.7	96	3	US-08-431-184-103	Sequence 103, App	500	5	3.7	116	2	US-08-082-842A-85	Sequence 85, Appl
428	5	3.7	97	3	US-08-300-928C-10	Sequence 10, Appl	501	5	3.7	116	2	US-08-672-345C-17	Sequence 17, App
429	5	3.7	97	3	US-08-430-944D-10	Sequence 10, Appl	502	5	3.7	116	4	US-08-839-765-167	Sequence 167, App
430	5	3.7	97	3	US-08-430-014-10	Sequence 10, Appl	503	5	3.7	116	4	US-08-839-765-168	Sequence 168, App
431	5	3.7	97	3	US-08-431-184-10	Sequence 10, Appl	504	5	3.7	116	4	US-09-136-389-167	Sequence 167, App
432	5	3.7	100	4	US-08-851-843A-10	Sequence 10, Appl	505	5	3.7	116	4	US-09-136-389-168	Sequence 168, App
433	5	3.7	100	4	US-08-974-549A-192	Sequence 192, App	506	5	3.7	116	4	US-08-484-537-3	Sequence 3, Appl
434	5	3.7	100	4	US-08-854-050-10	Sequence 10, Appl	507	5	3.7	116	4	US-08-484-537-14	Sequence 14, Appl
435	5	3.7	102	2	US-08-911-319A-4	Sequence 4, Appl	508	5	3.7	116	4	US-09-127-902-4	Sequence 4, Appl
436	5	3.7	102	3	US-09-352-619-4	Sequence 4, Appl	509	5	3.7	116	4	US-09-155-107-4	Sequence 4, Appl
437	5	3.7	107	4	US-08-654-482-6	Sequence 6, Appl	510	5	3.7	116	4	US-09-214-095D-17	Sequence 17, Appl
438	5	3.7	108	1	US-08-436-463-15	Sequence 15, Appl	511	5	3.7	116	4	PCT-US95-09641-4	Sequence 4, Appl
439	5	3.7	108	2	US-08-672-345C-103	Sequence 103, App	512	5	3.7	117	2	US-08-290-592E-17	Sequence 17, Appl
440	5	3.7	108	2	US-08-888-366-6	Sequence 6, Appl	513	5	3.7	117	2	US-08-290-592E-18	Sequence 18, Appl
441	5	3.7	108	4	US-09-214-095D-110	Sequence 110, App	514	5	3.7	117	2	US-08-672-345C-13	Sequence 13, Appl
442	5	3.7	109	1	US-07-841-646-21	Sequence 21, Appl	515	5	3.7	117	2	US-08-814-806-4	Sequence 4, Appl
443	5	3.7	109	1	US-07-662-193-5	Sequence 5, Appl	516	5	3.7	117	4	US-09-214-095D-13	Sequence 13, Appl
444	5	3.7	109	1	US-07-901-703-17	Sequence 17, Appl	517	5	3.7	117	5	PCT-US95-10053-14	Sequence 14, Appl
445	5	3.7	109	1	US-08-147-023-21	Sequence 21, Appl	518	5	3.7	117	5	PCT-US95-10053-15	Sequence 15, Appl
446	5	3.7	109	1	US-07-807-529A-6	Sequence 6, Appl	519	5	3.7	117	5	PCT-US96-09448-17	Sequence 17, Appl
447	5	3.7	109	1	US-08-447-570-21	Sequence 21, Appl	520	5	3.7	117	5	PCT-US96-09448-18	Sequence 18, Appl
448	5	3.7	109	2	US-08-210-762E-22	Sequence 22, Appl	521	5	3.7	117	6	5514582-7	Patent No. 5514582
449	5	3.7	109	2	US-08-449-700-21	Sequence 21, Appl	522	5	3.7	118	2	US-03-116-247-10	Sequence 10, Appl
450	5	3.7	109	2	US-08-449-699A-21	Sequence 21, Appl	523	5	3.7	119	1	US-08-192-102-5	Sequence 5, Appl
451	5	3.7	109	3	US-08-300-928C-8	Sequence 8, Appl	524	5	3.7	119	1	US-08-324-799-5	Sequence 5, Appl
452	5	3.7	109	3	US-09-053-197A-16	Sequence 16, Appl	525	5	3.7	119	1	US-08-497-312-20	Sequence 20, Appl
453	5	3.7	109	3	US-08-430-944D-8	Sequence 8, Appl	526	5	3.7	119	1	US-08-667-769A-15	Sequence 15, Appl
454	5	3.7	109	3	US-08-430-014-8	Sequence 8, Appl	527	5	3.7	119	1	US-03-340-536B-12	Sequence 12, Appl
455	5	3.7	109	3	US-08-431-184-8	Sequence 8, Appl	528	5	3.7	119	1	US-03-340-536A-13	Sequence 13, Appl
456	5	3.7	109	3	US-09-202-181-2	Sequence 2, Appl	529	5	3.7	119	2	US-03-303-569B-11	Sequence 11, Appl
457	5	3.7	109	4	PCT-US93-02462-6	Sequence 6, Appl	530	5	3.7	119	2	US-03-303-569B-12	Sequence 12, Appl
458	5	3.7	109	5	PCT-US93-05446-17	Sequence 17, Appl	531	5	3.7	119	2	US-03-303-569B-13	Sequence 13, Appl
459	5	3.7	111	1	US-07-662-193-4	Sequence 4, Appl	532	5	3.7	119	2	US-08-303-569B-14	Sequence 14, Appl
460	5	3.7	111	1	US-08-467-420A-15	Sequence 15, Appl	533	5	3.7	119	2	US-08-303-569B-15	Sequence 15, Appl
461	5	3.7	111	1	US-08-470-110A-15	Sequence 15, Appl	534	5	3.7	119	2	US-08-303-569B-16	Sequence 16, Appl
462	5	3.7	111	1	US-08-369-796-15	Sequence 15, Appl	535	5	3.7	119	2	US-08-303-569B-17	Sequence 17, Appl
463	5	3.7	111	2	US-08-940-371-15	Sequence 15, Appl	536	5	3.7	119	2	US-08-303-569B-18	Sequence 18, Appl
464	5	3.7	111	2	US-08-852-091-15	Sequence 15, Appl	537	5	3.7	119	2	US-08-303-569B-19	Sequence 19, Appl
465	5	3.7	111	3	US-08-300-928C-6	Sequence 6, Appl	538	5	3.7	119	2	US-08-303-569B-20	Sequence 20, Appl



539	5	3.7	119	2	US-08-303-569B-21	Sequence 21, Appl	612	5	3.7	134	2	US-08-436-717-39	Sequence 39, Appl
540	5	3.7	119	2	US-08-303-569B-22	Sequence 22, Appl	613	5	3.7	134	2	US-08-822-028-10	Sequence 10, Appl
541	5	3.7	119	2	US-08-303-569B-23	Sequence 23, Appl	614	5	3.7	134	4	US-08-479-285-10	Sequence 10, Appl
542	5	3.7	119	2	US-08-303-569B-24	Sequence 24, Appl	615	5	3.7	135	1	US-08-446-600A-4	Sequence 4, Appl
543	5	3.7	119	2	US-08-303-569B-30	Sequence 30, Appl	616	5	3.7	135	4	US-08-579-378A-16	Sequence 16, Appl
544	5	3.7	119	2	US-08-070-116A-10	Sequence 10, Appl	617	5	3.7	136	3	US-09-184-658-32	Sequence 32, Appl
545	5	3.7	119	2	US-08-192-861A-5	Sequence 5, Appl	618	5	3.7	136	3	US-08-649-100-25	Sequence 25, Appl
546	5	3.7	119	2	US-08-116-247-12	Sequence 12, Appl	619	5	3.7	137	2	US-08-379-057-18	Sequence 18, Appl
547	5	3.7	119	2	US-08-116-247-13	Sequence 13, Appl	620	5	3.7	137	2	US-08-621-751A-4	Sequence 1, Appl
548	5	3.7	119	2	US-08-116-247-14	Sequence 14, Appl	621	5	3.7	137	2	US-08-640-978C-1	Sequence 1, Appl
549	5	3.7	119	2	US-08-116-247-15	Sequence 15, Appl	622	5	3.7	137	2	US-08-640-978C-5	Sequence 5, Appl
550	5	3.7	119	2	US-08-116-247-16	Sequence 16, Appl	623	5	3.7	137	2	US-08-836-561-31	Sequence 31, Appl
551	5	3.7	119	2	US-08-116-247-17	Sequence 17, Appl	624	5	3.7	137	4	US-09-091-117-4	Sequence 4, Appl
552	5	3.7	119	2	US-08-116-247-18	Sequence 18, Appl	625	5	3.7	139	1	US-08-330-978-2	Sequence 2, Appl
553	5	3.7	119	2	US-08-116-247-19	Sequence 19, Appl	626	5	3.7	139	1	US-08-474-042-2	Sequence 2, Appl
554	5	3.7	119	2	US-08-116-247-20	Sequence 20, Appl	627	5	3.7	139	1	US-08-484-558-2	Sequence 2, Appl
555	5	3.7	119	2	US-08-116-247-21	Sequence 21, Appl	628	5	3.7	139	1	US-08-774-592-2	Sequence 2, Appl
556	5	3.7	119	2	US-08-116-247-22	Sequence 22, Appl	629	5	3.7	140	1	US-07-946-421-24	Sequence 24, Appl
557	5	3.7	119	2	US-08-116-247-23	Sequence 23, Appl	630	5	3.7	140	2	US-08-448-287-4	Sequence 2, Appl
558	5	3.7	119	2	US-08-116-247-24	Sequence 24, Appl	631	5	3.7	140	3	US-08-836-561-27	Sequence 27, Appl
559	5	3.7	119	2	US-08-116-247-25	Sequence 25, Appl	632	5	3.7	140	4	US-08-569-147-76	Sequence 76, Appl
560	5	3.7	119	2	US-08-767-128-18	Sequence 18, Appl	633	5	3.7	141	4	US-09-286-529-22	Sequence 22, Appl
561	5	3.7	119	4	US-09-133-119-5	Sequence 5, Appl	634	5	3.7	142	4	US-08-975-080-34	Sequence 34, Appl
562	5	3.7	119	4	US-08-192-093A-5	Sequence 5, Appl	635	5	3.7	143	1	US-08-236-520-7	Sequence 7, Appl
563	5	3.7	119	5	PCT-US95-17082A-15	Sequence 15, Appl	636	5	3.7	143	5	PCT-US95-05262-7	Sequence 7, Appl
564	5	3.7	120	1	US-08-111-080-25	Sequence 25, Appl	637	5	3.7	145	3	US-08-630-112-5	Sequence 5, Appl
565	5	3.7	120	1	US-08-211-980-25	Sequence 25, Appl	638	5	3.7	145	4	US-09-375-419-5	Sequence 5, Appl
566	5	3.7	120	1	US-08-497-312-26	Sequence 26, Appl	639	5	3.7	147	1	US-08-259-372A-8	Sequence 8, Appl
567	5	3.7	120	2	US-08-002-324-2	Sequence 2, Appl	640	5	3.7	147	1	US-08-468-671-8	Sequence 8, Appl
568	5	3.7	120	3	US-08-554-840-5	Sequence 5, Appl	641	5	3.7	149	3	US-08-584-031-17	Sequence 17, Appl
569	5	3.7	120	3	US-08-554-840-6	Sequence 6, Appl	642	5	3.7	150	2	US-08-400-115-2	Sequence 2, Appl
570	5	3.7	120	3	US-08-554-840-7	Sequence 7, Appl	643	5	3.7	150	2	US-08-403-866-8	Sequence 8, Appl
571	5	3.7	120	3	US-08-554-840-8	Sequence 8, Appl	644	5	3.7	158	1	US-08-836-075A-66	Sequence 66, Appl
572	5	3.7	120	4	US-09-188-930-180	Sequence 180, Appl	645	5	3.7	158	4	US-09-267-177-10	Sequence 10, Appl
573	5	3.7	120	5	PCT-US93-07967-25	Sequence 25, Appl	646	5	3.7	161	3	US-08-493-071-17	Sequence 17, Appl
574	5	3.7	120	5	PCT-US94-00261-2	Sequence 2, Appl	647	5	3.7	161	3	US-08-493-071-20	Sequence 20, Appl
575	5	3.7	121	1	US-08-339-582-2	Sequence 2, Appl	648	5	3.7	163	5	PCT-US91-02942-5	Sequence 5, Appl
576	5	3.7	121	1	US-08-974-899-4	Sequence 4, Appl	649	5	3.7	164	2	US-08-911-319A-1	Sequence 1, Appl
577	5	3.7	122	1	US-08-236-520-9	Sequence 9, Appl	650	5	3.7	164	3	US-09-352-619-1	Sequence 1, Appl
578	5	3.7	122	1	US-08-335-573-7	Sequence 7, Appl	651	5	3.7	168	1	US-08-460-739-5	Sequence 5, Appl
579	5	3.7	122	5	PCT-US95-05262-9	Sequence 9, Appl	652	5	3.7	171	2	US-08-417-495-24	Sequence 24, Appl
580	5	3.7	123	1	US-08-131-625B-17	Sequence 17, Appl	653	5	3.7	171	2	US-08-284-391B-24	Sequence 24, Appl
581	5	3.7	123	1	US-08-497-312-15	Sequence 15, Appl	654	5	3.7	171	3	US-08-463-903-20	Sequence 20, Appl
582	5	3.7	123	1	US-08-497-312-17	Sequence 17, Appl	655	5	3.7	171	4	US-09-218-950-24	Sequence 24, Appl
583	5	3.7	123	2	US-08-560-358E-28	Sequence 28, Appl	656	5	3.7	171	5	PCT-US92-0185-24	Sequence 24, Appl
584	5	3.7	123	2	US-08-799-464A-13	Sequence 13, Appl	657	5	3.7	171	5	PCT-US95-00454-24	Sequence 24, Appl
585	5	3.7	123	4	US-08-686-968C-11	Sequence 11, Appl	658	5	3.7	173	3	US-08-669-408B-8	Sequence 8, Appl
586	5	3.7	123	4	US-09-113-750A-42	Sequence 42, Appl	659	5	3.7	177	2	US-08-465-794-17	Sequence 17, Appl
587	5	3.7	123	5	PCT-US95-09927-13	Sequence 13, Appl	660	5	3.7	177	3	US-09-049-813-17	Sequence 17, Appl
588	5	3.7	123	5	PCT-US95-10904-19	Sequence 19, Appl	661	5	3.7	178	2	US-08-465-794-18	Sequence 18, Appl
589	5	3.7	123	5	PCT-US95-10904-53	Sequence 53, Appl	662	5	3.7	178	3	US-09-049-813-18	Sequence 18, Appl
590	5	3.7	123	5	PCT-US95-10904-55	Sequence 55, Appl	663	5	3.7	178	4	US-08-663-191A-4	Sequence 4, Appl
591	5	3.7	123	5	PCT-US95-10904-57	Sequence 57, Appl	664	5	3.7	179	3	US-08-649-100-9	Sequence 9, Appl
592	5	3.7	123	5	PCT-US95-10904-59	Sequence 59, Appl	665	5	3.7	179	4	US-08-836-500A-4	Sequence 4, Appl
593	5	3.7	123	5	PCT-US95-10904-61	Sequence 61, Appl	666	5	3.7	183	1	US-08-447-010-3	Sequence 3, Appl
594	5	3.7	124	4	US-08-751-359-14	Sequence 14, Appl	667	5	3.7	185	1	US-08-044-621D-36	Sequence 36, Appl
595	5	3.7	125	4	US-08-751-359-13	Sequence 13, Appl	668	5	3.7	185	1	US-08-044-621D-37	Sequence 37, Appl
596	5	3.7	126	2	US-08-244-205-15	Sequence 15, Appl	669	5	3.7	185	1	US-08-709-912-3	Sequence 3, Appl
597	5	3.7	126	4	US-08-751-359-11	Sequence 11, Appl	670	5	3.7	185	1	US-08-709-912-5	Sequence 5, Appl
598	5	3.7	126	4	US-08-751-359-12	Sequence 12, Appl	671	5	3.7	185	2	US-09-047-370-3	Sequence 3, Appl
599	5	3.7	126	5	PCT-US92-10284-15	Sequence 15, Appl	672	5	3.7	185	2	US-09-047-370-5	Sequence 5, Appl
600	5	3.7	130	3	US-08-836-662-4	Sequence 4, Appl	673	5	3.7	187	2	US-08-530-369B-19	Sequence 19, Appl
601	5	3.7	130	3	US-08-895-914-4	Sequence 4, Appl	674	5	3.7	187	3	US-08-493-071-16	Sequence 16, Appl
602	5	3.7	130	4	US-09-357-710A-4	Sequence 4, Appl	675	5	3.7	187	3	US-08-493-071-19	Sequence 19, Appl
603	5	3.7	132	4	US-08-434-000A-14	Sequence 14, Appl	676	5	3.7	188	2	US-08-484-993B-20	Sequence 20, Appl
604	5	3.7	132	4	US-09-312-157-14	Sequence 14, Appl	677	5	3.7	188	2	US-08-484-158B-20	Sequence 20, Appl
605	5	3.7	133	2	US-08-822-028-6	Sequence 6, Appl	678	5	3.7	188	2	US-08-484-156B-20	Sequence 20, Appl
606	5	3.7	133	2	US-08-822-028-30	Sequence 30, Appl	679	5	3.7	188	2	US-08-480-150A-20	Sequence 20, Appl
607	5	3.7	133	3	US-08-463-903-4	Sequence 4, Appl	680	5	3.7	188	3	US-08-458-731-20	Sequence 20, Appl
608	5	3.7	133	4	US-08-479-285-6	Sequence 6, Appl	681	5	3.7	188	3	US-08-149-223A-20	Sequence 20, Appl
609	5	3.7	133	4	US-08-479-285-30	Sequence 30, Appl	682	5	3.7	190	3	US-08-881-094-20	Sequence 20, Appl
610	5	3.7	133	6	5219996-17	Patent No. 5219996	683	5	3.7	191	2	US-08-290-665A-173	Sequence 173, App
611	5	3.7	134	1	US-08-137-117D-39	Sequence 39, Appl	684	5	3.7	191	5	PCT-US95-10398-173	Sequence 173, App

685	5	3.7	192	3	US-09-103-359-2	Sequence 2, Appl1	758	5	3.7	232	3	US-09-106-468-8	Sequence 8, Appl1
686	5	3.7	194	3	US-08-881-094-17	Sequence 17, Appl1	759	5	3.7	232	4	US-09-106-466A-8	Sequence 8, Appl1
687	5	3.7	196	4	US-09-383-586-11	Sequence 11, Appl1	760	5	3.7	232	4	US-09-106-467-8	Sequence 8, Appl1
688	5	3.7	200	1	US-07-855-412B-1	Sequence 1, Appl1	761	5	3.7	233	2	US-09-151-611-1	Sequence 1, Appl1
689	5	3.7	200	2	US-08-308-887A-1	Sequence 1, Appl1	762	5	3.7	233	4	US-09-370-102-1	Sequence 1, Appl1
690	5	3.7	200	3	US-08-881-094-1	Sequence 1, Appl1	763	5	3.7	234	4	US-08-944-483-46	Sequence 1, Appl1
691	5	3.7	200	3	US-08-881-094-19	Sequence 18, Appl1	764	5	3.7	235	2	US-08-716-301-10	Sequence 46, Appl1
692	5	3.7	202	3	US-08-881-094-19	Sequence 19, Appl1	765	5	3.7	236	3	US-08-493-071-15	Sequence 10, Appl1
693	5	3.7	208	1	US-08-315-695-21	Sequence 21, Appl1	766	5	3.7	236	3	US-08-493-071-15	Sequence 15, Appl1
694	5	3.7	210	4	US-08-961-083-36	Sequence 21, Appl1	767	5	3.7	237	2	US-08-469-537A-85	Sequence 18, Appl1
695	5	3.7	211	1	US-08-575-964-1	Sequence 36, Appl1	768	5	3.7	238	3	US-09-286-690-7	Sequence 85, Appl1
696	5	3.7	211	2	US-08-563-500-1	Sequence 1, Appl1	769	5	3.7	239	3	US-08-286-690-7	Sequence 7, Appl1
697	5	3.7	212	1	US-08-158-353-2	Sequence 1, Appl1	770	5	3.7	239	4	US-08-902-486-11	Sequence 8, Appl1
698	5	3.7	212	2	US-08-924-759-14	Sequence 2, Appl1	771	5	3.7	242	2	US-08-663-310-9	Sequence 11, Appl1
699	5	3.7	212	3	US-09-248-335-14	Sequence 14, Appl1	772	5	3.7	242	2	US-09-306-491-9	Sequence 9, Appl1
700	5	3.7	213	1	US-08-104-445-3	Sequence 3, Appl1	773	5	3.7	242	2	US-09-305-919-9	Sequence 9, Appl1
701	5	3.7	213	1	US-08-319-621A-12	Sequence 12, Appl1	774	5	3.7	244	4	US-08-980-832-29	Sequence 29, Appl1
702	5	3.7	213	1	US-08-165-038-4	Sequence 4, Appl1	775	5	3.7	245	3	US-08-415-788-33	Sequence 33, Appl1
703	5	3.7	213	1	US-08-621-081A-4	Sequence 4, Appl1	776	5	3.7	246	4	US-09-216-295-20	Sequence 2, Appl1
704	5	3.7	213	2	US-08-531-525-11	Sequence 11, Appl1	777	5	3.7	248	2	US-08-755-559-1	Sequence 20, Appl1
705	5	3.7	213	2	US-08-876-781-4	Sequence 4, Appl1	778	5	3.7	248	2	US-08-750-145A-18	Sequence 1, Appl1
706	5	3.7	213	2	US-08-718-270A-11	Sequence 11, Appl1	779	5	3.7	248	3	US-08-750-145A-20	Sequence 18, Appl1
707	5	3.7	217	1	US-08-621-081A-13	Sequence 13, Appl1	780	5	3.7	248	3	US-08-750-145A-20	Sequence 20, Appl1
708	5	3.7	217	1	US-08-621-081A-21	Sequence 21, Appl1	781	5	3.7	248	3	US-08-975-698A-22	Sequence 22, Appl1
709	5	3.7	217	1	US-08-621-081A-22	Sequence 22, Appl1	782	5	3.7	248	3	US-08-975-698A-24	Sequence 24, Appl1
710	5	3.7	217	1	US-08-621-081A-23	Sequence 23, Appl1	783	5	3.7	248	4	US-09-210-474-1	Sequence 1, Appl1
711	5	3.7	217	1	US-08-621-081A-24	Sequence 24, Appl1	784	5	3.7	248	4	US-09-417-090-22	Sequence 22, Appl1
712	5	3.7	217	1	US-08-621-081A-25	Sequence 25, Appl1	785	5	3.7	248	4	US-08-417-090-24	Sequence 24, Appl1
713	5	3.7	217	1	US-08-621-081A-26	Sequence 26, Appl1	786	5	3.7	249	2	US-08-155-111B-32	Sequence 32, Appl1
714	5	3.7	217	1	US-08-621-081A-27	Sequence 27, Appl1	787	5	3.7	249	2	US-08-433-998-32	Sequence 32, Appl1
715	5	3.7	217	1	US-08-621-081A-28	Sequence 28, Appl1	788	5	3.7	249	3	US-08-750-145A-3	Sequence 3, Appl1
716	5	3.7	217	1	US-08-621-081A-29	Sequence 29, Appl1	789	5	3.7	249	3	US-08-750-145A-10	Sequence 10, Appl1
717	5	3.7	217	1	US-08-621-081A-29	Sequence 29, Appl1	790	5	3.7	249	3	US-08-975-698A-7	Sequence 7, Appl1
718	5	3.7	219	1	PCT-US94-14106-59	Sequence 59, Appl1	791	5	3.7	249	4	US-08-591-468-2	Sequence 2, Appl1
719	5	3.7	219	1	US-09-247-373B-52	Sequence 52, Appl1	792	5	3.7	249	4	US-09-417-090-7	Sequence 7, Appl1
720	5	3.7	221	4	US-08-621-081A-19	Sequence 19, Appl1	793	5	3.7	249	4	US-09-417-090-7	Sequence 7, Appl1
721	5	3.7	222	1	US-07-688-352C-46	Sequence 46, Appl1	794	5	3.7	249	4	US-09-417-090-7	Sequence 7, Appl1
722	5	3.7	222	1	US-09-248-335-66	Sequence 66, Appl1	795	5	3.7	249	4	US-09-446-504-1	Sequence 4, Appl1
723	5	3.7	222	5	PCT-US81-02714-43	Sequence 43, Appl1	796	5	3.7	249	5	PCT-US94-06430-2	Sequence 2, Appl1
724	5	3.7	222	5	PCT-US91-02714-54	Sequence 54, Appl1	797	5	3.7	252	3	US-08-279-772A-6	Sequence 6, Appl1
725	5	3.7	223	4	US-09-120-426-4	Sequence 4, Appl1	798	5	3.7	252	4	US-08-902-486-9	Sequence 9, Appl1
726	5	3.7	223	5	PCT-US94-14106-51	Sequence 51, Appl1	799	5	3.7	253	2	US-08-474-379C-46	Sequence 46, Appl1
727	5	3.7	224	4	US-09-115-488-4	Sequence 4, Appl1	800	5	3.7	253	3	US-09-146-249A-46	Sequence 46, Appl1
728	5	3.7	224	4	US-08-448-398-1	Sequence 1, Appl1	801	5	3.7	253	3	US-08-206-188B-46	Sequence 46, Appl1
729	5	3.7	224	4	US-08-448-398-17	Sequence 17, Appl1	802	5	3.7	253	4	US-09-042-785A-4	Sequence 4, Appl1
730	5	3.7	224	4	US-08-447-703B-7	Sequence 7, Appl1	803	5	3.7	254	4	US-09-247-373B-38	Sequence 38, Appl1
731	5	3.7	225	4	US-09-247-373B-36	Sequence 36, Appl1	804	5	3.7	254	4	US-08-944-483-50	Sequence 50, Appl1
732	5	3.7	227	4	US-07-869-933-14	Sequence 14, Appl1	805	5	3.7	254	4	US-08-478-316-16	Sequence 16, Appl1
733	5	3.7	227	4	US-09-103-663-14	Sequence 14, Appl1	806	5	3.7	254	4	US-08-478-316-16	Sequence 16, Appl1
734	5	3.7	227	4	US-09-230-196-21	Sequence 21, Appl1	807	5	3.7	254	5	PCT-US95-10904-69	Sequence 69, Appl1
735	5	3.7	229	3	US-08-750-145A-4	Sequence 4, Appl1	808	5	3.7	255	5	US-09-027-337-7	Sequence 7, Appl1
736	5	3.7	229	3	US-08-975-698A-4	Sequence 4, Appl1	809	5	3.7	255	2	US-08-719-758-2	Sequence 2, Appl1
737	5	3.7	229	4	US-08-751-359-22	Sequence 22, Appl1	810	5	3.7	256	4	US-09-119-827-2	Sequence 4, Appl1
738	5	3.7	229	4	US-09-247-373B-48	Sequence 48, Appl1	811	5	3.7	261	1	US-07-940-605A-2	Sequence 2, Appl1
739	5	3.7	229	4	US-09-417-090-4	Sequence 4, Appl1	812	5	3.7	261	1	US-08-184-422-8	Sequence 8, Appl1
740	5	3.7	230	1	US-08-118-469A-3	Sequence 3, Appl1	813	5	3.7	261	1	US-08-360-923A-2	Sequence 2, Appl1
741	5	3.7	230	1	US-08-909-119-3	Sequence 3, Appl1	814	5	3.7	261	1	US-08-446-922-4	Sequence 4, Appl1
742	5	3.7	231	1	US-08-681-432-1	Sequence 1, Appl1	815	5	3.7	261	2	US-08-431-055-4	Sequence 4, Appl1
743	5	3.7	231	3	US-08-750-145A-11	Sequence 11, Appl1	816	5	3.7	261	2	US-08-690-995-12	Sequence 12, Appl1
744	5	3.7	231	3	US-08-975-698A-8	Sequence 8, Appl1	817	5	3.7	261	2	US-08-249-189-12	Sequence 12, Appl1
745	5	3.7	231	3	US-09-417-090-8	Sequence 8, Appl1	818	5	3.7	261	2	US-08-484-624A-12	Sequence 12, Appl1
746	5	3.7	232	1	US-08-278-091-8	Sequence 8, Appl1	819	5	3.7	261	2	US-08-477-733B-12	Sequence 12, Appl1
747	5	3.7	232	1	US-08-483-858-8	Sequence 8, Appl1	820	5	3.7	261	2	US-08-989-478-16	Sequence 16, Appl1
748	5	3.7	232	1	US-08-472-173-8	Sequence 8, Appl1	821	5	3.7	261	3	US-08-763-895-2	Sequence 3, Appl1
749	5	3.7	232	1	US-08-631-607-5	Sequence 8, Appl1	822	5	3.7	261	3	US-08-996-685-16	Sequence 16, Appl1
750	5	3.7	232	2	US-08-487-167-8	Sequence 8, Appl1	823	5	3.7	261	3	US-09-088-913A-12	Sequence 12, Appl1
751	5	3.7	232	2	US-08-482-816-8	Sequence 8, Appl1	824	5	3.7	261	3	US-08-589-371B-8	Sequence 8, Appl1
752	5	3.7	232	2	US-08-296-149-8	Sequence 8, Appl1	825	5	3.7	261	4	US-08-768-371-2	Sequence 2, Appl1
753	5	3.7	232	2	US-08-801-499-8	Sequence 8, Appl1	826	5	3.7	261	4	US-08-769-819-12	Sequence 12, Appl1
754	5	3.7	232	2	US-08-978-404B-45	Sequence 45, Appl1	827	5	3.7	261	4	US-08-770-974-12	Sequence 12, Appl1
755	5	3.7	232	2	US-08-615-271-8	Sequence 8, Appl1	828	5	3.7	261	5	US-08-858-197-4	Sequence 4, Appl1
756	5	3.7	232	2	US-09-074-660-8	Sequence 8, Appl1	829	5	3.7	261	5	PCT-US93-1003A-4	Sequence 4, Appl1
757	5	3.7	232	3	US-09-074-659-8	Sequence 8, Appl1	830	5	3.7	262	2	US-08-719-697-10	Sequence 10, Appl1

831	5	3.7	262	4	US-08-154-364-41	Sequence 41, Appli	904	5	3.7	296	1	US-08-428-927-2	Sequence 2, Appli
832	5	3.7	262	4	US-08-727-616A-10	Sequence 10, Appli	905	5	3.7	296	1	US-08-428-298-2	Sequence 2, Appli
833	5	3.7	263	4	US-08-776-059-43	Sequence 43, Appli	906	5	3.7	296	1	US-08-339-517-2	Sequence 2, Appli
834	5	3.7	264	1	US-08-107-042-2	Sequence 2, Appli	907	5	3.7	296	2	US-08-438-863-5	Sequence 5, Appli
835	5	3.7	264	2	US-07-857-224B-24	Sequence 24, Appli	908	5	3.7	296	3	US-08-438-862-5	Sequence 5, Appli
836	5	3.7	264	4	US-08-776-059-33	Sequence 33, Appli	909	5	3.7	297	2	US-08-504-459-12	Sequence 12, Appli
837	5	3.7	268	4	US-09-085-305-10	Sequence 10, Appli	910	5	3.7	301	2	US-08-656-906-25	Sequence 25, Appli
838	5	3.7	270	2	US-08-652-507-2	Sequence 2, Appli	911	5	3.7	301	4	US-09-217-847-25	Sequence 25, Appli
839	5	3.7	271	2	US-08-400-115-4	Sequence 4, Appli	912	5	3.7	306	1	US-08-424-682A-1	Sequence 1, Appli
840	5	3.7	271	4	US-08-957-351-26	Sequence 26, Appli	913	5	3.7	308	2	US-08-807-050-4	Sequence 4, Appli
841	5	3.7	273	1	US-08-446-922-11	Sequence 11, Appli	914	5	3.7	310	1	US-08-129-456A-37	Sequence 37, Appli
842	5	3.7	273	2	US-08-249-189-21	Sequence 21, Appli	915	5	3.7	311	1	US-07-701-935-26	Sequence 26, Appli
843	5	3.7	273	2	US-08-484-624A-21	Sequence 21, Appli	916	5	3.7	312	2	US-09-014-969-17	Sequence 17, Appli
844	5	3.7	273	3	US-08-477-733B-21	Sequence 21, Appli	917	5	3.7	312	4	US-08-360-821B-36	Sequence 36, Appli
845	5	3.7	273	2	US-09-088-913A-21	Sequence 21, Appli	918	5	3.7	315	2	US-08-910-927B-3	Sequence 3, Appli
846	5	3.7	273	4	US-08-769-819-21	Sequence 21, Appli	919	5	3.7	315	4	US-09-270-270-3	Sequence 2, Appli
847	5	3.7	273	4	US-08-770-974-21	Sequence 21, Appli	920	5	3.7	316	1	US-07-805-567-2	Sequence 2, Appli
848	5	3.7	275	3	US-08-463-903-6	Sequence 6, Appli	921	5	3.7	316	4	US-08-961-083-156	Sequence 156, App
849	5	3.7	277	1	US-08-690-457-5	Sequence 5, Appli	922	5	3.7	319	2	US-08-977-847-1	Sequence 1, Appli
850	5	3.7	277	2	US-08-628-187-5	Sequence 5, Appli	923	5	3.7	319	2	US-09-195-021-1	Sequence 1, Appli
851	5	3.7	277	3	US-08-493-071-3	Sequence 3, Appli	924	5	3.7	320	1	US-07-613-083B-1	Sequence 1, Appli
852	5	3.7	278	1	US-07-941-414-1	Sequence 1, Appli	925	5	3.7	320	2	US-08-933-750C-12	Sequence 12, Appli
853	5	3.7	278	1	US-08-249-377A-1	Sequence 1, Appli	926	5	3.7	320	4	US-09-234-613-12	Sequence 12, Appli
854	5	3.7	278	1	US-08-585-595-1	Sequence 1, Appli	927	5	3.7	323	4	US-09-237-543-6	Sequence 6, Appli
855	5	3.7	278	2	US-08-740-337-1	Sequence 1, Appli	928	5	3.7	323	6	5185254-4	Patent No. 5185254
856	5	3.7	278	3	US-09-184-658-47	Sequence 47, Appli	929	5	3.7	326	2	US-08-986-217-3	Sequence 3, Appli
857	5	3.7	278	3	US-09-260-283-2	Sequence 2, Appli	930	5	3.7	328	1	US-08-265-087-2	Sequence 2, Appli
858	5	3.7	279	1	US-08-690-457-11	Sequence 11, Appli	931	5	3.7	328	1	US-08-186-529-2	Sequence 2, Appli
859	5	3.7	279	2	US-08-628-187-1	Sequence 1, Appli	932	5	3.7	328	1	US-08-621-493-2	Sequence 2, Appli
860	5	3.7	279	3	US-08-493-071-6	Sequence 6, Appli	933	5	3.7	328	1	US-08-640-386A-2	Sequence 2, Appli
861	5	3.7	279	5	PCT-US95-00362-5	Sequence 5, Appli	934	5	3.7	328	2	US-08-965-688-2	Sequence 2, Appli
862	5	3.7	281	2	US-08-810-453-2	Sequence 2, Appli	935	5	3.7	328	2	US-08-751-767A-4	Sequence 4, Appli
863	5	3.7	281	4	US-08-815-190A-2	Sequence 2, Appli	936	5	3.7	328	4	US-09-260-173-2	Sequence 2, Appli
864	5	3.7	281	4	US-09-290-640-25	Sequence 25, Appli	937	5	3.7	331	1	US-08-364-081-3	Sequence 3, Appli
865	5	3.7	281	4	US-09-479-524-3	Sequence 3, Appli	938	5	3.7	331	5	US-08-630-552-3	Sequence 5, Appli
866	5	3.7	281	4	US-09-404-258-16	Sequence 16, Appli	939	5	3.7	331	5	PCT-US95-16558-3	Sequence 3, Appli
867	5	3.7	281	5	PCT-US95-00362-2	Sequence 2, Appli	940	5	3.7	334	1	US-08-287-442-9	Sequence 9, Appli
868	5	3.7	283	3	US-08-860-368B-21	Sequence 21, Appli	941	5	3.7	334	1	US-08-459-701-9	Sequence 9, Appli
869	5	3.7	283	4	US-08-961-083-124	Sequence 124, App	942	5	3.7	334	1	US-08-460-298-9	Sequence 9, Appli
870	5	3.7	284	3	US-09-053-197A-2	Sequence 2, Appli	943	5	3.7	334	1	US-08-459-174-9	Sequence 9, Appli
871	5	3.7	284	3	US-09-184-658-40	Sequence 40, Appli	944	5	3.7	334	4	US-08-711-417C-201	Sequence 201, App
872	5	3.7	284	4	US-08-793-701-39	Sequence 39, Appli	945	5	3.7	335	3	US-09-002-298-6	Sequence 6, Appli
873	5	3.7	284	4	US-08-793-701-41	Sequence 41, Appli	946	5	3.7	335	3	US-09-109-205-19	Sequence 19, Appli
874	5	3.7	284	4	US-08-793-701-57	Sequence 57, Appli	947	5	3.7	335	4	US-08-836-500A-2	Sequence 2, Appli
875	5	3.7	285	3	US-08-482-085B-20	Sequence 20, Appli	948	5	3.7	336	1	US-07-755-009-3	Sequence 3, Appli
876	5	3.7	285	3	US-08-463-903-22	Sequence 22, Appli	949	5	3.7	336	1	US-07-755-009-4	Sequence 4, Appli
877	5	3.7	285	4	US-09-095-855-197	Sequence 197, App	950	5	3.7	336	2	US-08-592-646A-1	Sequence 1, Appli
878	5	3.7	285	1	US-08-690-457-4	Sequence 4, Appli	951	5	3.7	336	2	US-08-474-379C-88	Sequence 88, Appli
879	5	3.7	287	2	US-08-628-187-4	Sequence 4, Appli	952	5	3.7	338	2	US-09-047-026A-2	Sequence 2, Appli
880	5	3.7	287	3	US-08-815-190A-16	Sequence 16, Appli	953	5	3.7	340	1	US-08-118-270-49	Sequence 49, Appli
881	5	3.7	287	3	US-08-493-071-2	Sequence 2, Appli	954	5	3.7	340	2	US-08-446-875-16	Sequence 16, Appli
882	5	3.7	287	4	US-09-097-767A-17	Sequence 17, Appli	955	5	3.7	340	4	US-09-360-779-2	Sequence 2, Appli
883	5	3.7	288	1	US-08-690-457-3	Sequence 3, Appli	956	5	3.7	340	5	PCT-US93-08528-49	Sequence 49, Appli
884	5	3.7	288	1	US-08-690-457-10	Sequence 10, Appli	957	5	3.7	343	3	US-08-446-875-2	Sequence 2, Appli
885	5	3.7	288	2	US-08-628-187-3	Sequence 3, Appli	958	5	3.7	343	2	US-09-123-851-1	Sequence 1, Appli
886	5	3.7	288	2	US-08-628-187-10	Sequence 10, Appli	959	5	3.7	343	2	US-08-102-385G-2	Sequence 2, Appli
887	5	3.7	288	3	US-08-493-071-1	Sequence 1, Appli	960	5	3.7	343	2	US-08-728-520-1	Sequence 1, Appli
888	5	3.7	288	3	US-08-493-071-5	Sequence 5, Appli	961	5	3.7	344	3	US-08-718-738-4	Sequence 4, Appli
889	5	3.7	289	1	US-08-690-457-9	Sequence 9, Appli	962	5	3.7	344	4	US-09-221-844-4	Sequence 4, Appli
890	5	3.7	289	2	US-08-628-187-9	Sequence 9, Appli	963	5	3.7	345	5	PCT-US95-03323A-4	Sequence 4, Appli
891	5	3.7	289	3	US-08-493-071-4	Sequence 4, Appli	964	5	3.7	346	2	US-08-744-779A-2	Sequence 2, Appli
892	5	3.7	289	4	US-09-077-675A-10	Sequence 10, Appli	965	5	3.7	346	2	US-08-864-531-2	Sequence 2, Appli
893	5	3.7	290	1	US-08-393-985-6	Sequence 6, Appli	966	5	3.7	347	1	US-07-637-870-5	Sequence 5, Appli
894	5	3.7	292	2	US-08-973-461A-4	Sequence 4, Appli	967	5	3.7	347	1	US-07-640-476-10	Sequence 10, Appli
895	5	3.7	292	3	US-08-648-010-4	Sequence 4, Appli	968	5	3.7	347	4	US-09-094-557-3	Sequence 3, Appli
896	5	3.7	293	2	US-08-489-141A-2	Sequence 2, Appli	969	5	3.7	348	1	US-07-637-399-8	Sequence 8, Appli
897	5	3.7	293	2	US-08-907-492A-4	Sequence 4, Appli	970	5	3.7	348	1	US-08-112-703-8	Sequence 8, Appli
898	5	3.7	293	2	US-08-576-626A-59	Sequence 59, Appli	971	5	3.7	348	2	US-08-974-546-1	Sequence 1, Appli
899	5	3.7	295	2	US-08-504-459-14	Sequence 14, Appli	972	5	3.7	348	3	US-08-852-824-17	Sequence 17, Appli
900	5	3.7	295	2	US-08-907-492A-2	Sequence 2, Appli	973	5	3.7	349	4	US-09-105-390-48	Sequence 48, Appli
901	5	3.7	296	1	US-08-428-926-2	Sequence 2, Appli	974	5	3.7	350	1	US-08-202-056-1	Sequence 1, Appli
902	5	3.7	296	1	US-08-435-434-5	Sequence 5, Appli	975	5	3.7	350	1	US-08-076-093A-2	Sequence 2, Appli
903	5	3.7	296	1	US-08-435-436-5	Sequence 5, Appli	976	5	3.7	350	1	US-08-450-393A-7	Sequence 7, Appli

977 5 3.7 350 1 US-08-410-453A-1 Sequence 1, Appl 1  
978 5 3.7 350 1 US-08-701-265-2 Sequence 2, Appl 1  
979 5 3.7 350 1 US-08-410-454A-1 Sequence 1, Appl 1  
980 5 3.7 350 2 US-08-284-586-2 Sequence 2, Appl 1  
981 5 3.7 350 2 US-08-410-456A-1 Sequence 2, Appl 1  
982 5 3.7 350 2 US-08-805-478-2 Sequence 2, Appl 1  
983 5 3.7 350 2 US-08-802-627A-2 Sequence 2, Appl 1  
984 5 3.7 350 2 US-08-801-238-2 Sequence 2, Appl 1  
985 5 3.7 350 2 US-08-801-228-2 Sequence 2, Appl 1  
986 5 3.7 350 2 US-08-966-116-16 Sequence 2, Appl 1  
987 5 3.7 350 4 US-09-104-296-2 Sequence 16, Appl 1  
988 5 3.7 350 4 US-08-446-669-7 Sequence 2, Appl 1  
989 5 3.7 350 5 PCT-US95-00476-7 Sequence 7, Appl 1  
990 5 3.7 321 1 US-08-248-466B-12 Sequence 12, Appl 1  
991 5 3.7 332 1 US-08-967-272-2 Sequence 2, Appl 1  
992 5 3.7 333 2 US-08-865-203-4 Sequence 4, Appl 1  
993 5 3.7 333 2 US-07-849-420-4 Sequence 4, Appl 1  
994 5 3.7 333 4 US-09-253-854-4 Sequence 4, Appl 1  
995 5 3.7 353 4 US-08-855-424-4 Sequence 4, Appl 1  
996 5 3.7 354 1 US-07-759-568-2 Sequence 2, Appl 1  
997 5 3.7 354 3 US-08-969-644-10 Sequence 10, Appl 1  
998 5 3.7 354 3 US-08-444-189-10 Sequence 2, Appl 1  
999 5 3.7 354 4 US-09-068-569-2 Sequence 10, Appl 1  
1000 5 3.7 354 4 US-08-468-544-10 Sequence 10, Appl 1

## ALIGNMENTS

RESULT 1  
US-08-371-505-2  
; Sequence 2, Application US/08371505  
; Patent No. 5695754  
; GENERAL INFORMATION:  
; APPLICANT: COLLEN, DESIRE  
; TITLE OF INVENTION: STAPHYLOKINASE DERIVATIVES  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WEBB ZIESENHEIM BRUENING LOGSDON ORKIN & HANSON, P.C.  
; STREET: 700 KOPPERS BUILDING, 436 SEVENTH AVENUE  
; CITY: PITTSBURGH  
; STATE: PENNSYLVANIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 15222-2363  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: NEC 286  
; OPERATING SYSTEM: DOS  
; SOFTWARE: WORDPERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/371,505  
; FILING DATE: 11 JAN 1995  
; CLASSIFICATION: 424  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 136  
; TYPE: AMINO ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: UNKNOWN  
US-08-371-505-2

Query Match 100.0%; Score 136; DB 1; Length 136;  
Best Local Similarity 100.0%; Pred. No. 4.7e-133;  
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSSFDKGYKKGGDDASYEPTGPIYLMVNTGVDSKGNELSPHYVEPIKPGTTLTKRKI 60  
DB 1 SSSFDKGYKKGGDDASYEPTGPIYLMVNTGVDSKGNELSPHYVEPIKPGTTLTKRKI 60  
OY 61 EYVEMALDATAVYKEFRVVELDPSAKIEVYTYDKNKKKEETKSPITEKGFVVDLSHI 120  
DB 61 EYVEMALDATAVYKEFRVVELDPSAKIEVYTYDKNKKKEETKSPITEKGFVVDLSHI 120

OY 121 KNPGFNLTIVYIEKK 136  
DB 121 KNPGFNLTIVYIEKK 136

RESULT 2  
US-08-784-971-5  
; Sequence 5, Application US/08784971  
; Patent No. 5951980  
; GENERAL INFORMATION:  
; APPLICANT: COLLEN, DESIRE  
; TITLE OF INVENTION: NEW STAPHYLOKINASE DERIVATIVES  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: THE WEBB LAW FIRM  
; STREET: 700 KOPPERS BUILDING, 436 SEVENTH AVENUE  
; CITY: PITTSBURGH  
; STATE: PENNSYLVANIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 15219-1818  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" FLOPPY DISK  
; COMPUTER: DIGITAL VENTURIS GL 6200  
; OPERATING SYSTEM: DOS  
; SOFTWARE: MICROSOFT WORD 2.0c  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/784,971  
; FILING DATE: 16-JAN-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/371,505  
; FILING DATE: 11-JAN-1995  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 136  
; TYPE: AMINO ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: UNKNOWN  
US-08-784-971-5

Query Match 100.0%; Score 136; DB 2; Length 136;  
Best Local Similarity 100.0%; Pred. No. 4.7e-133;  
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSSFDKGYKKGGDDASYEPTGPIYLMVNTGVDSKGNELSPHYVEPIKPGTTLTKRKI 60  
DB 1 SSSFDKGYKKGGDDASYEPTGPIYLMVNTGVDSKGNELSPHYVEPIKPGTTLTKRKI 60  
OY 61 EYVEMALDATAVYKEFRVVELDPSAKIEVYTYDKNKKKEETKSPITEKGFVVDLSHI 120  
DB 61 EYVEMALDATAVYKEFRVVELDPSAKIEVYTYDKNKKKEETKSPITEKGFVVDLSHI 120  
OY 121 KNPGFNLTIVYIEKK 136  
DB 121 KNPGFNLTIVYIEKK 136

RESULT 3  
US-08-256-261-6  
; Sequence 6, Application US/08256261  
; Patent No. 5801037  
; GENERAL INFORMATION:  
; APPLICANT: Behnke, Detlef  
; APPLICANT: Schlotz, Bernhard  
; APPLICANT: Albrecht, Sybille  
; APPLICANT: G hrs, Karl-Heinz  
; APPLICANT: Hartmann, Manfred  
; TITLE OF INVENTION: Expression of signal-peptide-free  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:

```
;; ADDRESSEE: Fish & Neave
;; STREET: 1251 Avenue of the Americas
;; CITY: New York
;; STATE: New York
;; COUNTRY: USA
;; ZIP: 10020
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentln Release #1.0, Version #1.30 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/256,261
;; FILING DATE:
;; CLASSIFICATION: 435
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 137 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-256-261-6
```

```
Query Match          100.0%; Score 136; DB 1; Length 137;
Best Local Similarity 100.0%; Pred. No. 4.8e-133;
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 SSSFDKGGYKKGDASYPEPTGPIYLMVNVTVGDSKGNELLSPHYVEFPKPGTTLTKRKI 60
    |||||||
DB 2 SSSFDKGGYKKGDASYPEPTGPIYLMVNVTVGDSKGNELLSPHYVEFPKPGTTLTKRKI 61
QY 61 EYVEMALDATAKKEFRVVELDPSAKIEVTVYDKNKKKEETKSPITEKGFPVPLSEHI 120
    |||||||
DB 62 EYVEMALDATAKKEFRVVELDPSAKIEVTVYDKNKKKEETKSPITEKGFPVPLSEHI 121
QY 121 KNPGFNLTIKVIEKK 136
    |||||||
DB 122 KNPGFNLTIKVIEKK 137
```

```
RESULT 4
US-08-852-299-6
;; Sequence 6, Application US/08852299
;; Patent No. 6010897
;; GENERAL INFORMATION:
;; APPLICANT: Behnke, Detlef
;; APPLICANT: Schloft, Bernhard
;; APPLICANT: Albrecht, Sybille
;; APPLICANT: G hrs, Karl-Heinz
;; APPLICANT: Hartmann, Manfred
;; TITLE OF INVENTION: Expression of signal-peptide-free
;; NUMBER OF SEQUENCES: 40
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Fish & Neave
;; STREET: 1251 Avenue of the Americas
;; CITY: New York
;; STATE: New York
;; COUNTRY: USA
;; ZIP: 10020
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentln Release #1.0, Version #1.30 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/852,299
;; FILING DATE: 17-MAY-1997
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/256,261
;; FILING DATE:
```

```
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 137 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-852-299-6
```

```
Query Match          100.0%; Score 136; DB 3; Length 137;
Best Local Similarity 100.0%; Pred. No. 4.8e-133;
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 SSSFDKGGYKKGDASYPEPTGPIYLMVNVTVGDSKGNELLSPHYVEFPKPGTTLTKRKI 60
    |||||||
DB 2 SSSFDKGGYKKGDASYPEPTGPIYLMVNVTVGDSKGNELLSPHYVEFPKPGTTLTKRKI 61
QY 61 EYVEMALDATAKKEFRVVELDPSAKIEVTVYDKNKKKEETKSPITEKGFPVPLSEHI 120
    |||||||
DB 62 EYVEMALDATAKKEFRVVELDPSAKIEVTVYDKNKKKEETKSPITEKGFPVPLSEHI 121
QY 121 KNPGFNLTIKVIEKK 136
    |||||||
DB 122 KNPGFNLTIKVIEKK 137
```

```
RESULT 5
US-08-256-261-4
;; Sequence 4, Application US/08256261
;; Patent No. 5801037
;; GENERAL INFORMATION:
;; APPLICANT: Behnke, Detlef
;; APPLICANT: Schloft, Bernhard
;; APPLICANT: Albrecht, Sybille
;; APPLICANT: G hrs, Karl-Heinz
;; APPLICANT: Hartmann, Manfred
;; TITLE OF INVENTION: Expression of signal-peptide-free
;; NUMBER OF SEQUENCES: 40
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Fish & Neave
;; STREET: 1251 Avenue of the Americas
;; CITY: New York
;; STATE: New York
;; COUNTRY: USA
;; ZIP: 10020
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentln Release #1.0, Version #1.30 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/256,261
;; FILING DATE:
;; CLASSIFICATION: 435
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 137 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-256-261-4
```

```
Query Match          75.0%; Score 102; DB 1; Length 137;
Best Local Similarity 100.0%; Pred. No. 6.1e-98;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 35 KGNELLSPHYVEFPKPGTTLTKRKIEYVEMALDATAKKEFRVVELDPSAKIEVTVYDK 94
    |||||||
DB 36 KGNELLSPHYVEFPKPGTTLTKRKIEYVEMALDATAKKEFRVVELDPSAKIEVTVYDK 95
QY 95 NKKKEETKSPITEKGFPVPLSEHIKNPGFNLTIKVIEKK 136
```

Db 96 NKKKEETKSPITEKGFVVPDLSEHIKNPGFNLTIKVIEKK 137

RESULT 6  
US-08-852-299-4  
Sequence 4, Application US/08852299

Patent No. 6010897

GENERAL INFORMATION:

APPLICANT: Behnke, Detlef

APPLICANT: Schloft, Bernhard

APPLICANT: Albrecht, Sybille

APPLICANT: G hrs, Karl-Heinz

APPLICANT: Hartmann, Manfred

TITLE OF INVENTION: Expression of signal-peptide-free

TITLE OF INVENTION: staphylokinases

NUMBER OF SEQUENCES: 40

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Neave

STREET: 1251 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10020

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/852,299

FILING DATE: 17-MAY-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/256,261

FILING DATE:

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 137 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-852-299-4

Query Match 75.0%; Score 102; DB 3; Length 137;  
Best Local Similarity 100.0%; Pred. No. 6.1e-98;  
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 KKNELISPHYVEPIKPGTTLTKETIEYVEMALDATAKKEFRVVELDPSAKIEVTYDK 94

Db 36 KKNELISPHYVEPIKPGTTLTKETIEYVEMALDATAKKEFRVVELDPSAKIEVTYDK 95

QY 95 NKKKEETKSPITEKGFVVPDLSEHIKNPGFNLTIKVIEKK 136

Db 96 NKKKEETKSPITEKGFVVPDLSEHIKNPGFNLTIKVIEKK 137

RESULT 7  
US-08-256-261-10  
Sequence 10, Application US/08256261

Patent No. 5801037

GENERAL INFORMATION:

APPLICANT: Behnke, Detlef

APPLICANT: Schloft, Bernhard

APPLICANT: Albrecht, Sybille

APPLICANT: G hrs, Karl-Heinz

APPLICANT: Hartmann, Manfred

TITLE OF INVENTION: Expression of signal-peptide-free

TITLE OF INVENTION: staphylokinases

NUMBER OF SEQUENCES: 40

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Neave

STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/256,261

FILING DATE:

CLASSIFICATION: 435

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 123 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-256-261-10

Query Match 68.4%; Score 93; DB 1; Length 123;  
Best Local Similarity 100.0%; Pred. No. 1.1e-88;  
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 YVEPIKPGTTLTKETIEYVEMALDATAKKEFRVVELDPSAKIEVTYDKNKKKEETKS 103

Db 31 YVEPIKPGTTLTKETIEYVEMALDATAKKEFRVVELDPSAKIEVTYDKNKKKEETKS 90

QY 104 FPIKEKGFVVPDLSEHIKNPGFNLTIKVIEKK 136

Db 91 FPIKEKGFVVPDLSEHIKNPGFNLTIKVIEKK 123

RESULT 8  
US-08-852-299-10

Sequence 10, Application US/08852299

Patent No. 6010897

GENERAL INFORMATION:

APPLICANT: Behnke, Detlef

APPLICANT: Schloft, Bernhard

APPLICANT: Albrecht, Sybille

APPLICANT: G hrs, Karl-Heinz

APPLICANT: Hartmann, Manfred

TITLE OF INVENTION: Expression of signal-peptide-free

TITLE OF INVENTION: staphylokinases

NUMBER OF SEQUENCES: 40

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Neave

STREET: 1251 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10020

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/852,299

FILING DATE: 17-MAY-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/256,261

FILING DATE:

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 123 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-852-299-10

Query Match  
Best Local Similarity 68.4%; Score 93; DB 3; Length 123;  
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 YVEFPKPGTTLTKKEIEYVEMALDATAKKEFRVVELDPSAKIEVYTDKNNKKEETKS 103  
DB 31 YVEFPKPGTTLTKKEIEYVEMALDATAKKEFRVVELDPSAKIEVYTDKNNKKEETKS 90

QY 104 FPITEKGFVVPDLSEHIKNPGFNLTIKVIEKK 136  
DB 91 FPITEKGFVVPDLSEHIKNPGFNLTIKVIEKK 123

RESULT 9  
US-08-256-261-8  
; Sequence 8, Application US/08256261  
; Patent No. 5801037  
; GENERAL INFORMATION:  
; APPLICANT: Behnke, Detlef  
; APPLICANT: Schlotz, Bernhard  
; APPLICANT: Albrecht, Sybille  
; APPLICANT: G hrs, Karl-Heinz  
; APPLICANT: Hartmann, Manfred  
; TITLE OF INVENTION: Expression of signal-peptide-free  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Neave  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/256,261  
; FILING DATE:  
; CLASSIFICATION: 435  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 127 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-256-261-8

Query Match  
Best Local Similarity 68.4%; Score 93; DB 1; Length 127;  
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 YVEFPKPGTTLTKKEIEYVEMALDATAKKEFRVVELDPSAKIEVYTDKNNKKEETKS 103  
DB 35 YVEFPKPGTTLTKKEIEYVEMALDATAKKEFRVVELDPSAKIEVYTDKNNKKEETKS 94

QY 104 FPITEKGFVVPDLSEHIKNPGFNLTIKVIEKK 136  
DB 95 FPITEKGFVVPDLSEHIKNPGFNLTIKVIEKK 127

RESULT 10  
US-08-852-299-8  
; Sequence 8, Application US/08852299  
; Patent No. 6010897  
; GENERAL INFORMATION:

APPLICANT: Behnke, Detlef  
APPLICANT: Schlotz, Bernhard  
APPLICANT: Albrecht, Sybille  
APPLICANT: G hrs, Karl-Heinz  
APPLICANT: Hartmann, Manfred  
TITLE OF INVENTION: Expression of signal-peptide-free  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Neave  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/852,299  
FILING DATE: 17-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/256,261  
FILING DATE:  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 127 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-852-299-8

Query Match  
Best Local Similarity 68.4%; Score 93; DB 3; Length 127;  
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 YVEFPKPGTTLTKKEIEYVEMALDATAKKEFRVVELDPSAKIEVYTDKNNKKEETKS 103  
DB 35 YVEFPKPGTTLTKKEIEYVEMALDATAKKEFRVVELDPSAKIEVYTDKNNKKEETKS 94

QY 104 FPITEKGFVVPDLSEHIKNPGFNLTIKVIEKK 136  
DB 95 FPITEKGFVVPDLSEHIKNPGFNLTIKVIEKK 127

RESULT 11  
US-08-256-261-2  
; Sequence 2, Application US/08256261  
; Patent No. 5801037  
; GENERAL INFORMATION:  
; APPLICANT: Behnke, Detlef  
; APPLICANT: Schlotz, Bernhard  
; APPLICANT: Albrecht, Sybille  
; APPLICANT: G hrs, Karl-Heinz  
; APPLICANT: Hartmann, Manfred  
; TITLE OF INVENTION: Expression of signal-peptide-free  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Neave  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,261  
FILING DATE:  
CLASSIFICATION: 435  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 137 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-256-261-2

Query Match 68.4%; Score 93; DB 1; Length 137;  
Best Local Similarity 100.0%; Pred. No. 1.2e-88;  
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 YVEFPKPGTTLTKKEIEYVEMALDATAVKEFRVVELDPSAKIEVTYDKNKKKEETKS 103  
|||||  
DB 45 YVEFPKPGTTLTKKEIEYVEMALDATAVKEFRVVELDPSAKIEVTYDKNKKKEETKS 104  
|||||  
QY 104 FPITEKGFVVPDLSEHKNGFNLTITKVIEKK 136  
|||||  
DB 105 FPITEKGFVVPDLSEHKNGFNLTITKVIEKK 137  
|||||

RESULT 12  
US-08-256-261-12  
Sequence 12, Application US/08256261  
Patent No. 5801037  
GENERAL INFORMATION:  
APPLICANT: Behnke, Detlef  
APPLICANT: Schloft, Bernhard  
APPLICANT: Albrecht, Sybille  
APPLICANT: G hrs, Karl-Heinz  
APPLICANT: Hartmann, Manfred  
TITLE OF INVENTION: Expression of signal-peptide-free  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Fish & Neave  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,261  
FILING DATE:  
CLASSIFICATION: 435  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 137 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-256-261-12

Query Match 68.4%; Score 93; DB 1; Length 137;  
Best Local Similarity 100.0%; Pred. No. 1.2e-88;  
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 YVEFPKPGTTLTKKEIEYVEMALDATAVKEFRVVELDPSAKIEVTYDKNKKKEETKS 103  
|||||  
DB 45 YVEFPKPGTTLTKKEIEYVEMALDATAVKEFRVVELDPSAKIEVTYDKNKKKEETKS 104  
|||||

QY 104 FPITEKGFVVPDLSEHKNGFNLTITKVIEKK 136  
|||||  
DB 105 FPITEKGFVVPDLSEHKNGFNLTITKVIEKK 137  
|||||

RESULT 13  
US-08-256-261-14  
Sequence 14, Application US/08256261  
Patent No. 5801037  
GENERAL INFORMATION:  
APPLICANT: Behnke, Detlef  
APPLICANT: Schloft, Bernhard  
APPLICANT: Albrecht, Sybille  
APPLICANT: G hrs, Karl-Heinz  
APPLICANT: Hartmann, Manfred  
TITLE OF INVENTION: Expression of signal-peptide-free  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Fish & Neave  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,261  
FILING DATE:  
CLASSIFICATION: 435  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 137 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-256-261-14

Query Match 68.4%; Score 93; DB 1; Length 137;  
Best Local Similarity 100.0%; Pred. No. 1.2e-88;  
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 YVEFPKPGTTLTKKEIEYVEMALDATAVKEFRVVELDPSAKIEVTYDKNKKKEETKS 103  
|||||  
DB 45 YVEFPKPGTTLTKKEIEYVEMALDATAVKEFRVVELDPSAKIEVTYDKNKKKEETKS 104  
|||||  
QY 104 FPITEKGFVVPDLSEHKNGFNLTITKVIEKK 136  
|||||  
DB 105 FPITEKGFVVPDLSEHKNGFNLTITKVIEKK 137  
|||||

RESULT 14  
US-08-852-299-2  
Sequence 2, Application US/08852299  
Patent No. 6010897  
GENERAL INFORMATION:  
APPLICANT: Behnke, Detlef  
APPLICANT: Schloft, Bernhard  
APPLICANT: Albrecht, Sybille  
APPLICANT: G hrs, Karl-Heinz  
APPLICANT: Hartmann, Manfred  
TITLE OF INVENTION: Expression of signal-peptide-free  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Fish & Neave  
STREET: 1251 Avenue of the Americas  
CITY: New York



STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/852,299  
FILING DATE: 17-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/256,261  
FILING DATE:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 137 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-852-299-2

Query Match 68.4%; Score 93; DB 3; Length 137;  
Best Local Similarity 100.0%; Pred. No. 1.2e-88;  
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 YVEFPKPGTTLTKKEIEYVEMALDATAYKEFRVVELDPSAKIEVYTDKNNKKEETKS 103  
|||||  
DB 45 YVEFPKPGTTLTKKEIEYVEMALDATAYKEFRVVELDPSAKIEVYTDKNNKKEETKS 104

QY 104 PPITEKGFVVDLSEHIKNGFNLTITKVIEKK 136  
|||||  
DB 105 PPITEKGFVVDLSEHIKNGFNLTITKVIEKK 137

RESULT 15  
US-08-852-299-12  
Sequence 12, Application US/08852299  
Patent No. 6010897  
GENERAL INFORMATION:  
APPLICANT: Behnke, Detlef  
APPLICANT: Schlotz, Bernhard  
APPLICANT: Albrecht, Sybille  
APPLICANT: G hrs, Karl-Heinz  
APPLICANT: Hartmann, Manfred  
TITLE OF INVENTION: Expression of signal-peptide-free  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Neave  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/852,299  
FILING DATE: 17-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/256,261  
FILING DATE:  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 137 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-852-299-12

Query Match 68.4%; Score 93; DB 3; Length 137;  
Best Local Similarity 100.0%; Pred. No. 1.2e-88;  
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 YVEFPKPGTTLTKKEIEYVEMALDATAYKEFRVVELDPSAKIEVYTDKNNKKEETKS 103  
|||||  
DB 45 YVEFPKPGTTLTKKEIEYVEMALDATAYKEFRVVELDPSAKIEVYTDKNNKKEETKS 104

QY 104 PPITEKGFVVDLSEHIKNGFNLTITKVIEKK 136  
|||||  
DB 105 PPITEKGFVVDLSEHIKNGFNLTITKVIEKK 137

Search completed: April 22, 2002, 10:45:41  
Job time: 116 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 10:44:00 ; Search time 14.45 Seconds

(without alignments)  
716.937 Million cell updates/sec

Title: US-09-601-490-1

Perfect score: 136  
Sequence: 1 SSSFPDKGKYKKGGDASYFPP.....SEHIKNPGFNLTIKVIEKK 136

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size: 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database:

PIR\_68:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	102	75.0	163	1 PRSAK	staphylokinase - p
2	93	68.4	163	2 S02330	staphylokinase - p
3	8	5.9	245	3 A72218	gufa protein - The
4	7	5.1	56	2 G69044	hypothetical prote
5	7	5.1	90	2 I46953	CGMP-gated retinal
6	7	5.1	92	2 PQ0629	coat protein - tul
7	7	5.1	102	2 C84026	ribosomal protein
8	7	5.1	138	2 DB3704	transposase (14) B
9	7	5.1	194	2 E71501	hypothetical prote
10	7	5.1	204	2 DB4190	tryptophan synthas
11	7	5.1	219	2 T27336	hypothetical prote
12	7	5.1	277	2 A56044	tryptophan synthas
13	7	5.1	288	2 S21499	coat protein - tur
14	7	5.1	288	2 S51329	polypotein - tur
15	7	5.1	311	2 A83746	transposase (16) B
16	7	5.1	334	2 T47256	thioredoxin reduct
17	7	5.1	337	2 A48347	coat protein - tur
18	7	5.1	343	2 E72694	hypothetical prote
19	7	5.1	371	2 C83736	transposase (15) B
20	7	5.1	388	2 PQ0757	polypotein - tur
21	7	5.1	388	2 PQ0759	polypotein - tur
22	7	5.1	398	2 T72335	hypothetical prote
23	7	5.1	399	2 T32933	hypothetical prote
24	7	5.1	408	2 S11689	coat protein - tur
25	7	5.1	424	2 T29158	hypothetical prote
26	7	5.1	694	2 A75127	hypothetical prote
27	7	5.1	722	2 S64492	hypothetical prote
28	7	5.1	944	2 H64650	translation initia
29	7	5.1	1008	2 T18508	hypothetical prote

30	7	5.1	1032	2 S12153	pol polypotein -
31	7	5.1	1034	1 GNLJCA	pol polypotein -
32	7	5.1	1035	1 GNLJGG	pol polypotein -
33	7	5.1	1036	1 GNLJG2	pol polypotein -
34	7	5.1	1055	1 GNLJST	pol polypotein -
35	7	5.1	1055	2 S53092	pol polypotein -
36	7	5.1	3163	1 U01895	genome polypotein
37	6	4.4	63	2 G69941	hypothetical prote
38	6	4.4	70	2 A31940	microtubule-associ
39	6	4.4	77	2 H82017	pseudogene (IS1655
40	6	4.4	85	2 S44604	C02F5.2 protein -
41	6	4.4	89	2 T45510	probable transport
42	6	4.4	90	2 PQ0292	polypotein - gior
43	6	4.4	93	2 A56623	hypothetical prote
44	6	4.4	129	1 A56623	glycine cleavage s
45	6	4.4	129	2 C85946	hypothetical prote
46	6	4.4	138	1 RKKKSC	ribulose-bisphosph
47	6	4.4	138	2 D72420	hypothetical prote
48	6	4.4	141	2 H86609	kinase [imported]
49	6	4.4	146	2 E83190	(38)-hydroxymyrist
50	6	4.4	146	2 T33016	hypothetical prote
51	6	4.4	153	1 S08619	formate hydrogenly
52	6	4.4	153	2 F43868	hypothetical prote
53	6	4.4	153	2 D85921	hypothetical prote
54	6	4.4	158	2 S78249	ribosomal protein
55	6	4.4	158	2 F83861	hypothetical prote
56	6	4.4	160	2 D70364	ribosomal protein
57	6	4.4	160	2 G70457	ribosomal protein
58	6	4.4	162	2 S13195	ganglioside M2 act
59	6	4.4	162	2 F86690	hypothetical prote
60	6	4.4	165	2 T52114	probable transcrip
61	6	4.4	166	2 B70369	cytochrome c552 -
62	6	4.4	166	2 G85057	hypothetical prote
63	6	4.4	173	2 JC5610	toponin I - sea s
64	6	4.4	175	2 A46388	Her-1 protein - Ca
65	6	4.4	175	2 DB3304	conserved hypotet
66	6	4.4	183	1 C69474	conserved hypotet
67	6	4.4	187	2 G70226	conserved hypotet
68	6	4.4	191	2 H84472	hypothetical prote
69	6	4.4	193	2 I54178	ganglioside M2 act
70	6	4.4	199	2 G81261	probable lipoprote
71	6	4.4	200	2 S22411	ganglioside M2 act
72	6	4.4	200	2 D86645	hypothetical prote
73	6	4.4	201	2 E84187	hypothetical prote
74	6	4.4	210	2 F70546	hypothetical prote
75	6	4.4	215	2 T23482	hypothetical prote
76	6	4.4	216	2 T29842	hypothetical prote
77	6	4.4	216	2 T34527	hypothetical prote
78	6	4.4	219	2 A35650	Sur protein - chic
79	6	4.4	221	1 A26753	glutathione trans
80	6	4.4	223	1 A41031	glutathione trans
81	6	4.4	226	2 DB8875	hypothetical prote
82	6	4.4	227	2 A71224	probable fibrillar
83	6	4.4	227	2 H75191	fibrillar-like p
84	6	4.4	227	2 T41133	hypothetical prote
85	6	4.4	230	2 A40085	ubiquitin thiolest
86	6	4.4	230	2 JC7117	ubiquitin carboxy-
87	6	4.4	230	2 T34854	hypothetical prote
88	6	4.4	231	2 F70219	thiol-disulfide in
89	6	4.4	231	2 E81215	hypothetical prote
90	6	4.4	232	2 T29841	hypothetical prote
91	6	4.4	234	1 A70549	probable methyltra
92	6	4.4	237	2 C84968	cell division inh
93	6	4.4	240	2 S36797	lectin BMA - Bowri
94	6	4.4	241	2 B69655	two-component resp
95	6	4.4	244	2 T48110	hypothetical prote
96	6	4.4	248	2 E64898	H repeat-associate
97	6	4.4	252	2 E75491	time protein - Del
98	6	4.4	254	2 D82124	vacu lipoprotein V
99	6	4.4	265	2 S60947	hypothetical prote
100	6	4.4	269	1 A27067	calretinin - chick
101	6	4.4	269	2 F85063	hypothetical prote
102	6	4.4	270	2 T06992	translation initia

pol polypotein -  
pol polypotein -  
pol polypotein -  
pol polypotein -  
pol polypotein -  
pol polypotein -  
genome polypotein  
hypothetical prote  
microtubule-associ  
pseudogene (IS1655  
C02F5.2 protein -  
probable transport  
polypotein - gior  
hypothetical prote  
glycine cleavage s  
hypothetical prote  
ribulose-bisphosph  
hypothetical prote  
kinase [imported]  
(38)-hydroxymyrist  
hypothetical prote  
formate hydrogenly  
hypothetical prote  
hypothetical prote  
ribosomal protein  
hypothetical prote  
ribosomal protein  
ribosomal protein  
ganglioside M2 act  
hypothetical prote  
probable transcrip  
cytochrome c552 -  
hypothetical prote  
toponin I - sea s  
Her-1 protein - Ca  
conserved hypotet  
conserved hypotet  
conserved hypotet  
hypothetical prote  
ganglioside M2 act  
probable lipoprote  
ganglioside M2 act  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
Sur protein - chic  
glutathione trans  
glutathione trans  
hypothetical prote  
probable fibrillar  
fibrillar-like p  
hypothetical prote  
ubiquitin thiolest  
ubiquitin carboxy-  
hypothetical prote  
thiol-disulfide in  
hypothetical prote  
probable methyltra  
cell division inh  
lectin BMA - Bowri  
two-component resp  
hypothetical prote  
H repeat-associate  
time protein - Del  
vacu lipoprotein V  
hypothetical prote  
calretinin - chick  
hypothetical prote  
translation initia

103	6	4.4	270	2	T35365	hypothetical prote
104	6	4.4	271	1	JX0316	phosphatidylinosit
105	6	4.4	272	2	T19537	hypothetical prote
106	6	4.4	274	2	T05864	hypothetical prote
107	6	4.4	275	2	E83655	signal peptidase-1
108	6	4.4	275	2	D70737	hypothetical prote
109	6	4.4	275	2	G85536	SKM dependent meth
110	6	4.4	275	2	D72066	sam dependent meth
111	6	4.4	276	2	T46368	hypothetical prote
112	6	4.4	277	2	G71849	amino acid ABC tra
113	6	4.4	279	2	A82922	conserved hypotet
114	6	4.4	279	2	F70612	hypothetical prote
115	6	4.4	279	2	C70101	hypothetical prote
116	6	4.4	281	2	H64125	modb protein homol
117	6	4.4	282	2	T31941	hypothetical prote
118	6	4.4	283	2	S10773	2-hydroxymuconic s
119	6	4.4	285	2	C45691	main capsid protei
120	6	4.4	292	2	F71901	hypothetical prote
121	6	4.4	292	2	E64614	beta-alanine synth
122	6	4.4	295	2	T34774	hypothetical prote
123	6	4.4	297	2	E69152	formylmethanofuran
124	6	4.4	299	2	H71576	probable alpha(1,2
125	6	4.4	300	2	A83653	acetylcholinester
126	6	4.4	301	2	S62087	hrpf protein - pse
127	6	4.4	302	2	G84053	hybrid-endo-beta-1
128	6	4.4	303	2	A25095	nodulation protei
129	6	4.4	304	2	T23919	hypothetical prote
130	6	4.4	313	2	S26838	nodulation protei
131	6	4.4	316	2	E82266	trehalase operon r
132	6	4.4	320	1	A43824	periplasmic flagel
133	6	4.4	321	2	G86400	hypothetical prote
134	6	4.4	321	2	E83839	hypothetical prote
135	6	4.4	332	2	A83857	3-dehydroquinatase
136	6	4.4	333	2	A48957	tumor surface anti
137	6	4.4	334	2	T25521	hypothetical prote
138	6	4.4	337	1	S75716	molibdenum cofacto
139	6	4.4	337	1	S15303	probable CDP-6-deo
140	6	4.4	331	2	T22648	hypothetical prote
141	6	4.4	332	2	T32863	hypothetical prote
142	6	4.4	340	2	S69194	N4-(beta-N-acetylgl
143	6	4.4	344	2	A75456	conserved hypotet
144	6	4.4	344	2	PC4185	hemagglutinin 1 ch
145	6	4.4	345	2	T37576	probable mitochond
146	6	4.4	345	2	S36080	hemagglutinin - In
147	6	4.4	345	2	JQ1308	hemagglutinin 1 -
148	6	4.4	345	2	JQ1309	hemagglutinin 1 -
149	6	4.4	345	2	JQ1313	hemagglutinin 1 -
150	6	4.4	345	2	JQ1306	hemagglutinin 1 -
151	6	4.4	345	2	JQ1315	hemagglutinin 1 -
152	6	4.4	345	2	JQ1316	hemagglutinin 1 -
153	6	4.4	345	2	JQ1311	hemagglutinin 1 -
154	6	4.4	345	2	JQ1314	hemagglutinin 1 -
155	6	4.4	345	2	JQ1312	hemagglutinin 1 -
156	6	4.4	345	2	JQ1307	hemagglutinin 1 -
157	6	4.4	345	2	JQ1310	hemagglutinin 1 -
158	6	4.4	347	2	JQ1301	hemagglutinin 1 -
159	6	4.4	347	2	JQ1304	hemagglutinin 1 -
160	6	4.4	347	2	JQ1302	hemagglutinin 1 -
161	6	4.4	347	2	JQ1305	hemagglutinin 1 -
162	6	4.4	347	2	JQ1303	hemagglutinin 1 -
163	6	4.4	348	2	A83532	hypothetical prote
164	6	4.4	349	2	D70383	hydrogenase (EC 1.
165	6	4.4	350	2	A75600	conserved hypotet
166	6	4.4	353	1	HMI3A	hemagglutinin prec
167	6	4.4	361	2	T12571	cinnaamyl-alcohol d
168	6	4.4	361	2	H71202	hypothetical prote
169	6	4.4	365	2	T08577	hypothetical prote
170	6	4.4	367	2	G84236	hypothetical prote
171	6	4.4	368	2	C70786	probable ilvE prot
172	6	4.4	373	2	G84840	hypothetical prote
173	6	4.4	373	2	S42434	hypothetical prote
174	6	4.4	375	2	B40205	Na+/H+-exchanging
175	6	4.4	376	2	E84742	hypothetical prote
176	6	4.4	376	2	T17128	hypothetical prote
177	6	4.4	377	1	S01615	site-specific DNA-
178	6	4.4	377	2	T40024	probable cytochrom
179	6	4.4	377	2	T21751	hypothetical prote
180	6	4.4	378	2	S11758	hemagglutinin prec
181	6	4.4	378	2	T18486	hypothetical prote
182	6	4.4	383	2	D75288	carbohydrate kinas
183	6	4.4	390	2	C85510	hypothetical prote
184	6	4.4	392	2	S04718	DNA-directed RNA p
185	6	4.4	392	2	A44167	aminomethyltransfe
186	6	4.4	396	2	S31151	translation elonga
187	6	4.4	396	2	E84186	cell division prot
188	6	4.4	396	2	A57080	CSA protein - huma
189	6	4.4	397	2	JC5385	translational elonga
190	6	4.4	398	2	JC4254	ribosomal protein
191	6	4.4	398	2	T15846	interleukin-1 rece
192	6	4.4	407	2	S17428	hypothetical prote
193	6	4.4	407	2	G84783	probable pectinest
194	6	4.4	407	2	F70451	hypothetical prote
195	6	4.4	409	2	F72504	probable 26S prote
196	6	4.4	410	2	A44391	serum response ele
197	6	4.4	412	2	G64685	hypothetical prote
198	6	4.4	412	2	S48881	rRNA (guanosine-2'
199	6	4.4	413	2	A69205	sensory transducti
200	6	4.4	414	1	C70859	probable hexosyltr
201	6	4.4	418	2	E96687	hypothetical prote
202	6	4.4	421	2	H72492	probable proton/so
203	6	4.4	424	2	T08412	hypothetical prote
204	6	4.4	424	2	E83384	probable ring-hydr
205	6	4.4	424	2	T31978	hypothetical prote
206	6	4.4	425	2	D75130	amino acid transpo
207	6	4.4	425	2	F71075	probable protein k
208	6	4.4	426	2	H84349	peroxidase / catal
209	6	4.4	429	2	S61186	glutamate 5'-kinase
210	6	4.4	429	2	T33832	hypothetical prote
211	6	4.4	431	2	G82430	conserved hypotet
212	6	4.4	431	2	F69458	coenzyme F390 synt
213	6	4.4	434	2	D71309	47K integral membr
214	6	4.4	434	2	G84046	hypothetical prote
215	6	4.4	435	2	T27039	hypothetical prote
216	6	4.4	436	2	B70157	hypothetical prote
217	6	4.4	437	2	A64891	coenzyme F390 synt
218	6	4.4	438	2	T22060	hypothetical prote
219	6	4.4	443	2	G83350	hypothetical prote
220	6	4.4	446	2	T31644	probable carboxyl-
221	6	4.4	447	2	G75383	hypothetical prote
222	6	4.4	450	2	G83740	hypothetical prote
223	6	4.4	452	2	I39383	UDP-N-acetylglucos
224	6	4.4	453	2	D82036	ntic protein - Rho
225	6	4.4	455	2	S43175	regulatory protein
226	6	4.4	457	2	A54604	1-aminocyclopropan
227	6	4.4	463	2	S27757	glycosyltransferas
228	6	4.4	465	2	S56176	probable glucosylt
229	6	4.4	466	2	T00090	hypothetical prote
230	6	4.4	466	2	T22141	hypothetical prote
231	6	4.4	466	2	F84414	hypothetical prote
232	6	4.4	469	2	A64411	hypothetical prote
233	6	4.4	476	2	T03747	glucosyltransferas
234	6	4.4	476	2	B82414	leucine aminopepti
235	6	4.4	476	2	S60381	REP3 protein homol
236	6	4.4	480	1	JH0531	d phosphogluconate d
237	6	4.4	481	1	T28900	hypothetical prote
238	6	4.4	481	2	S28429	alanine transaminas
239	6	4.4	482	2	S70887	phosphomannomutase
240	6	4.4	483	2	T44335	hypothetical prote
241	6	4.4	483	2	T48119	hypothetical prote
242	6	4.4	486	2	B69680	pata-nitrobenzyl e
243	6	4.4	489	2	C71410	hypothetical prote
244	6	4.4	495	2	S50508	AMP1 protein - yea
245	6	4.4	500	2	S23243	hypothetical prote
246	6	4.4	503	2	T05272	fatty acid elongas
247	6	4.4	506	2		
248	6	4.4				

249	6	4.4	508	2	E71620	hypothetical prote	322	6	4.4	748	1	S08680	methionyl-CoA
250	6	4.4	508	2	T20757	hypothetical prote	323	6	4.4	750	1	A59145	methionyl-CoA
251	6	4.4	511	2	T51544	hypothetical prote	324	6	4.4	753	1	T32844	hypothetical prote
252	6	4.4	513	2	T45658	1-phosphatidylinos	325	6	4.4	756	2	G84866	hypothetical prote
253	6	4.4	520	2	T45765	hypothetical prote	326	6	4.4	757	2	D84175	zinc-transporing
254	6	4.4	520	2	T94978	probable glutamate	327	6	4.4	760	2	F75530	ribonuclease - Dei
255	6	4.4	522	2	C96608	hypothetical prote	328	6	4.4	764	2	E84516	probable retroelem
256	6	4.4	523	2	D83631	probable sulfate t	329	6	4.4	767	2	T00360	hypothetical prote
257	6	4.4	526	2	A34896	adenylate cyclase-	330	6	4.4	768	2	F75350	probable protein-e
258	6	4.4	529	2	T48467	aspartyl aminopept	331	6	4.4	769	1	P18PF6	pI protein - phage
259	6	4.4	531	1	S44649	f4h10.7 protein -	332	6	4.4	771	2	T34376	hypothetical prote
260	6	4.4	531	1	S54098	1-phosphatidylinos	333	6	4.4	772	2	C75579	catalase - Deinoco
261	6	4.4	532	2	T06029	hypothetical prote	334	6	4.4	772	2	T02805	chloride channel p
262	6	4.4	534	2	S26877	groEL protein - re	335	6	4.4	782	2	T22134	hypothetical prote
263	6	4.4	535	2	T52098	probable nuclear t	336	6	4.4	794	2	S50687	hypothetical prote
264	6	4.4	537	2	A86154	hypothetical prote	337	6	4.4	794	2	F72202	hypothetical prote
265	6	4.4	538	2	B81376	probable membrane	338	6	4.4	798	2	S62791	probable lipoprote
266	6	4.4	540	2	E69861	ABC transporter (A	339	6	4.4	800	2	S53079	PER111 protein - Y
267	6	4.4	548	2	C75499	groEL protein - De	340	6	4.4	803	2	B84931	DNA topoisomerase
268	6	4.4	550	2	G86667	hypothetical prote	341	6	4.4	807	2	T32463	hypothetical prote
269	6	4.4	550	2	T29919	hypothetical prote	342	6	4.4	814	1	C40618	fibrial outer mem
270	6	4.4	553	2	H64494	hypothetical prote	343	6	4.4	821	1	S67087	hypothetical prote
271	6	4.4	565	2	G82443	conserved hypotet	344	6	4.4	822	2	H69547	molybdopterin oxid
272	6	4.4	569	2	T43531	probable potassium	345	6	4.4	822	2	I48742	ryanodine receptor
273	6	4.4	574	1	HMIIBM	hemagglutinin prec	346	6	4.4	827	1	A36895	endopeptidase Ia (
274	6	4.4	575	1	HMIIBM	hemagglutinin prec	347	6	4.4	827	1	S56404	virulence associat
275	6	4.4	575	1	S01862	hemagglutinin prec	348	6	4.4	827	1	C86114	probable enzyme va
276	6	4.4	576	2	S03300	hemagglutinin prec	349	6	4.4	837	2	T48407	hypothetical prote
277	6	4.4	578	2	S03299	hemagglutinin prec	350	6	4.4	837	2	JN0292	antigen 332 - mala
278	6	4.4	580	2	S03301	hemagglutinin prec	351	6	4.4	839	2	C84697	hypothetical prote
279	6	4.4	580	2	A82447	probable 5'-nucleo	352	6	4.4	844	2	T32608	hypothetical prote
280	6	4.4	581	2	T38501	hypothetical prote	353	6	4.4	856	2	H64552	endopeptidase Clp
281	6	4.4	583	1	HMIIBS	hemagglutinin prec	354	6	4.4	862	2	T05941	lipoxigenase (EC 1
282	6	4.4	583	1	HMIIBS	hemagglutinin prec	355	6	4.4	862	2	T40934	probable integral
283	6	4.4	584	1	HMIIB	hemagglutinin prec	356	6	4.4	865	2	T11852	lipoxigenase (EC 1
284	6	4.4	585	1	HMIIBJ	hemagglutinin prec	357	6	4.4	867	2	B96625	hypothetical prote
285	6	4.4	610	2	H71612	asparagine--tRNA l	358	6	4.4	874	2	B70945	hypothetical prote
286	6	4.4	614	2	A96716	probable fructokin	359	6	4.4	875	2	I59350	hypothetical prote
287	6	4.4	627	2	S46820	hypothetical prote	360	6	4.4	876	2	T05943	probable lipoxigen
288	6	4.4	628	1	A39262	transcription fact	361	6	4.4	876	2	I52907	importin beta chai
289	6	4.4	628	1	A33333	transcription fact	362	6	4.4	876	2	S66288	nuclear pore-targe
290	6	4.4	630	1	A46149	transcription fact	363	6	4.4	887	2	T20941	hypothetical prote
291	6	4.4	631	2	T13115	protein gp29 - pha	364	6	4.4	892	2	S76350	hypothetical prote
292	6	4.4	632	2	T38126	probable electron	365	6	4.4	902	2	S61144	glycogen phosphory
293	6	4.4	635	2	S57714	cspB protein - Cio	366	6	4.4	902	2	B84652	hypothetical prote
294	6	4.4	638	1	TVXLRF	protein kinase raf	367	6	4.4	906	2	A71438	probable resistanc
295	6	4.4	640	2	D83571	conserved hypotet	368	6	4.4	910	2	S07838	SEC15 protein - ye
296	6	4.4	644	2	C86747	topoisomerase IV s	369	6	4.4	949	2	E71940	translation initia
297	6	4.4	651	2	F64457	methionine--tRNA l	370	6	4.4	961	2	T01167	hypothetical prote
298	6	4.4	651	2	A36333	dnak-type molecula	371	6	4.4	966	1	PHPOAG	starch phosphoryla
299	6	4.4	652	2	D85044	hypothetical prote	372	6	4.4	969	2	T15446	hypothetical prote
300	6	4.4	662	2	D54078	methy1-accepting c	373	6	4.4	984	2	E70406	DMSO reductase cha
301	6	4.4	664	2	S66067	methionine--tRNA l	374	6	4.4	987	2	H86197	hypothetical prote
302	6	4.4	668	2	T15305	hypothetical prote	375	6	4.4	987	2	T40241	probable guanine n
303	6	4.4	672	2	T32557	hypothetical prote	376	6	4.4	997	2	S33754	glutamate receptor
304	6	4.4	676	2	A81349	probable ATP-depen	377	6	4.4	1000	2	I46521	titin - rabbit (fr
305	6	4.4	677	2	G69895	formate dehydrogen	378	6	4.4	1013	2	T16244	hypothetical prote
306	6	4.4	684	2	F85075	hypothetical prote	379	6	4.4	1018	1	F73720	cytodifference acces
307	6	4.4	686	2	A71607	Mtn3/RAG1P-like p	380	6	4.4	1025	2	S50293	multidrug resistanc
308	6	4.4	696	2	S75626	hypothetical prote	381	6	4.4	1025	2	T18376	hypothetical prote
309	6	4.4	700	2	T43469	hypothetical prote	382	6	4.4	1029	2	H96658	hypothetical prote
310	6	4.4	704	2	T34034	hypothetical prote	383	6	4.4	1036	2	B69368	hypothetical prote
311	6	4.4	705	2	A41322	N-acetylmuramoyl-L	384	6	4.4	1096	2	T48512	hypothetical prote
312	6	4.4	706	2	A45990	junctional sarcopl	385	6	4.4	1120	2	T38431	DNA-directed RNA p
313	6	4.4	710	2	H72341	vacu protein - The	386	6	4.4	1132	2	T03844	telomerase catalyt
314	6	4.4	715	2	S70397	zona pellucida gly	387	6	4.4	1211	2	C71817	DNA polymerase III
315	6	4.4	720	2	T36293	probable serine/th	388	6	4.4	1211	2	D64702	DNA polymerase III
316	6	4.4	728	1	A35644	hepatocyte growth	389	6	4.4	1225	1	B64234	hypothetical prote
317	6	4.4	728	1	A60185	hepatocyte growth	390	6	4.4	1235	2	C69165	hypothetical prote
318	6	4.4	732	2	T08420	1-phosphatidylinos	391	6	4.4	1243	2	UC5615	membrane-associate
319	6	4.4	737	2	T31349	hypothetical prote	392	6	4.4	1262	2	T22523	hypothetical prote
320	6	4.4	737	2	T15615	hypothetical prote	393	6	4.4	1270	2	T26720	hypothetical prote
321	6	4.4	744	2	A70385	DNA gyrase A subun	394	6	4.4	1272	2	T49313	copia-type reverse

395	4.4	1274	2	D84485	probable retroelem	468	5	3.7	54	2	E45681	hypothetical 5.9K
396	4.4	1276	2	T03711	probable calmoduli	469	5	3.7	55	1	BGRT	spermatid transiti
397	4.4	1284	2	G83897	conserved hypothet	470	5	3.7	55	1	BGMS	hypothetical prote
398	4.4	1300	2	T53799	CGI protein - huma	471	5	3.7	56	2	B86805	hypothetical prote
399	4.4	1320	2	F96614	probable copia-lyp	472	5	3.7	56	2	H81353	hypothetical prote
400	4.4	1352	2	F86246	hypothetical prote	473	5	3.7	57	2	D49056	T-cell receptor al
401	4.4	1352	2	T47925	copia-type polypro	474	5	3.7	58	2	A35416	DNA topoisomerase
402	4.4	1354	2	S74244	serine/threonine-s	475	5	3.7	59	2	C23727	neurotoxin V-5 - b
403	4.4	1354	2	S69212	serine/threonine-s	476	5	3.7	59	2	A72392	hypothetical prote
404	4.4	1356	2	S32763	kinectin 1 - huma	477	5	3.7	59	2	D81363	small hydrophobic
405	4.4	1369	2	T32338	hypothetical prote	478	5	3.7	59	2	C36493	H+-transporting AT
406	4.4	1466	2	G84516	probable retroelem	479	5	3.7	61	2	B33488	hypothetical prote
407	4.4	1486	2	C64832	cell division prot	480	5	3.7	61	2	S42858	hypothetical prote
408	4.4	1486	2	E85618	hypothetical prote	481	5	3.7	62	2	JQ2194	hypothetical 6.7K
409	4.4	1534	2	JH0228	cell division prot	482	5	3.7	64	2	PC4361	pepsinogen C-1 - h
410	4.4	1629	2	T06461	DNA-binding protei	483	5	3.7	65	2	T09827	polysialacturonase
411	4.4	1650	2	T18444	hypothetical prote	484	5	3.7	66	2	T07880	protein kinase (EC
412	4.4	1742	2	S76110	hypothetical prote	485	5	3.7	67	2	G69058	hypothetical prote
413	4.4	1787	2	F84528	probable retroelem	486	5	3.7	69	2	A41903	recombinase homolo
414	4.4	1872	2	JC4976	plexin 3 precursor	487	5	3.7	69	2	E86644	hypothetical prote
415	4.4	1883	2	G82875	hypothetical prote	488	5	3.7	69	2	T45696	hypothetical prote
416	4.4	1888	2	T14273	zinc finger protei	489	5	3.7	70	2	T06525	cytochrome P450 ho
417	4.4	1894	2	JC4980	plexin 1 precursor	490	5	3.7	70	2	T12789	bacteriocin homolo
418	4.4	1905	2	T53553	plexin - African c	491	5	3.7	71	1	IHC	high potential lro
419	4.4	1948	2	B69511	N conserved hypoch	492	5	3.7	74	2	S68269	Y-box binding prot
420	4.4	1957	2	S68453	sodium channel pro	493	5	3.7	75	2	A25155	Ig heavy chain V r
421	4.4	1968	1	S05697	myosin heavy chain	494	5	3.7	75	2	C70523	hypothetical prote
422	4.4	2172	2	T00936	probable ATP-depen	495	5	3.7	76	1	S76753	hypothetical prote
423	4.4	2206	2	G71611	hypothetical prote	496	5	3.7	76	1	S11704	aminoglycoside 3''
424	4.4	2232	2	T34434	hypothetical prote	497	5	3.7	76	2	F70317	RNA polymerase ome
425	4.4	2331	2	T25410	hypothetical prote	498	5	3.7	78	2	A82011	probable integral
426	4.4	2340	2	B71704	cell surface antiq	499	5	3.7	78	2	D72764	probable ferredoxi
427	4.4	2485	1	H71621	serine/threonine-s	500	5	3.7	79	2	B83230	probable acyl carr
428	4.4	3079	1	RC8Y12	probable GTPase-ac	501	5	3.7	79	2	G64375	hypothetical prote
429	4.4	3488	2	T34418	hypothetical prote	502	5	3.7	80	2	S25050	Ig heavy chain V r
430	4.4	3519	2	S43048	polyketide synthas	503	5	3.7	80	2	E64307	ferredoxin - Metha
431	4.4	4558	2	C82199	RTX toxin RtxA Vcl	504	5	3.7	81	2	C86119	hypothetical prote
432	4.4	4589	2	T14914	dynamin beta heavy	505	5	3.7	81	2	D65233	hypothetical 9.6 K
433	4.4	4688	2	F82885	hypothetical prote	506	5	3.7	81	2	B84149	hypothetical prote
434	4.4	4687	2	S72269	ryanodine receptor	507	5	3.7	82	2	I51133	MHC class II beta
435	4.4	4699	2	A37113	ryanodine receptor	508	5	3.7	82	2	S22306	hypothetical prote
436	4.4	4981	2	T18489	hypothetical prote	509	5	3.7	83	2	I46058	caldesmon - bovine
437	4.4	6805	2	S20901	titin - rabbit (fr	510	5	3.7	84	2	PH1487	Ig heavy chain V r
438	4.4	26926	1	I38344	titin, cardiac mus	511	5	3.7	85	2	T12867	hypothetical prote
439	3.7	17	2	S20490	photosystem II chl	512	5	3.7	86	2	T03624	reverse transcript
440	3.7	17	2	S66364	sodium-translocat	513	5	3.7	86	2	H83457	hypothetical prote
441	3.7	21	2	C49042	Ig heavy chain V r	514	5	3.7	87	2	T03716	reverse transcript
442	3.7	21	2	B49042	Ig heavy chain V r	515	5	3.7	87	2	T06281	hypothetical prote
443	3.7	25	2	PH1715	Ig heavy chain V r	516	5	3.7	87	2	B85510	hypothetical 10K p
444	3.7	26	2	T14041	NADH dehydrogenase	517	5	3.7	87	2	JQ1264	hypothetical prote
445	3.7	26	2	T14221	chromogranin-B - r	518	5	3.7	87	2	T23257	hypothetical prote
446	3.7	32	2	F49164	phospholipase A2 (	519	5	3.7	88	2	S27275	glutaredoxin 3 - S
447	3.7	32	2	S58413	hypothetical prote	520	5	3.7	88	2	S77429	Ig heavy chain V r
448	3.7	32	2	H82416	ribosomal protein	521	5	3.7	88	2	E25155	anti sigma factor
449	3.7	33	2	S22602	hypothetical prote	522	5	3.7	88	2	G70331	DNA-binding protei
450	3.7	38	2	T19769	hypothetical prote	523	5	3.7	89	2	G81106	reverse transcript
451	3.7	39	2	G83716	hypothetical prote	524	5	3.7	89	2	T14589	hypothetical prote
452	3.7	41	2	I44530	T-cell receptor al	525	5	3.7	89	2	D64414	hypothetical prote
453	3.7	42	2	S06812	Myosin heavy chain	526	5	3.7	89	2	B32227	meRE protein homol
454	3.7	42	2	S06813	Ig heavy chain V r	527	5	3.7	89	2	T44502	hypothetical prote
455	3.7	44	2	PL0091	hypothetical prote	528	5	3.7	89	2	T25923	ribosomal protein
456	3.7	44	2	E37286	olfactory receptor	529	5	3.7	91	1	R3MX10	phospholipase D (E
457	3.7	45	2	JC5274	thymosin beta - hu	530	5	3.7	91	2	PC4194	regulatory proteins
458	3.7	45	2	PL0094	Ig heavy chain V r	531	5	3.7	92	1	Z1BPC2	cytastatin protein
459	3.7	48	2	S61469	p83/100 protein -	532	5	3.7	92	2	JN0906	reverse transcript
460	3.7	48	2	S61472	p83/100 protein -	533	5	3.7	92	2	T03623	hypothetical prote
461	3.7	49	1	T0B01	thymopoietin I - b	534	5	3.7	92	2	S76032	hypothetical prote
462	3.7	49	1	T0B02	thymopoietin II -	535	5	3.7	92	2	S56289	heterocyst diftere
463	3.7	49	1	T0B03	thymopoietin III -	536	5	3.7	93	2	C24672	Ig heavy chain V r
464	3.7	52	1	RUDVEG	rubredoxin [valida	537	5	3.7	93	2	A25343	nonhistone chromos
465	3.7	52	2	A71002	hypothetical prote	538	5	3.7	93	2	S13717	histone-like prote
466	3.7	53	2	B25507	proteinase inhibit	539	5	3.7	93	2	F86063	hypothetical prote
467	3.7	54	2	F82711	50S ribosomal prot	540	5	3.7	93	2	S48658	peptidylprolyl iso

541	5	3.7	93	2	E69510	proteinase inhibit	614	5	3.7	107	2	JC1127	major allergen cha
542	5	3.7	94	2	G32513	Ig heavy chain V r	615	5	3.7	107	2	T47876	hypothetical prote
543	5	3.7	94	2	JC1164	hypothetical 10.3K	616	5	3.7	107	2	T10041	hypothetical prote
544	5	3.7	94	2	C59094	hypothetical prote	617	5	3.7	107	2	T42903	hypothetical prote
545	5	3.7	94	2	D75066	lsu ribosomal prot	618	5	3.7	107	2	T42918	hypothetical prote
546	5	3.7	95	2	S12888	DNA-binding protei	619	5	3.7	108	2	PI0248	Ig heavy chain V r
547	5	3.7	96	1	FEXCAL	ferredoxin [2Fe-2S	620	5	3.7	108	2	S31953	Ig heavy chain V r
548	5	3.7	96	2	S16215	xyulokinas (EC 2	621	5	3.7	108	2	S31954	penicillin-binding
549	5	3.7	96	2	JC5945	regulatory protein	622	5	3.7	108	2	S77095	hypothetical prote
550	5	3.7	96	2	S38916	hypothetical prote	623	5	3.7	108	2	C64510	hypothetical prote
551	5	3.7	96	2	S78713	protein YDR322c-a	624	5	3.7	108	2	T26681	hypothetical prote
552	5	3.7	97	1	FEXCT	ferredoxin [2Fe-2S	625	5	3.7	109	1	L1HWVA	Ig lambda chain V-
553	5	3.7	97	2	T15891	hypothetical prote	626	5	3.7	109	1	R3MX17	ribosomal protein
554	5	3.7	98	1	KRG1BS	keratin, feather -	627	5	3.7	109	2	S27053	thioredoxin - Emer
555	5	3.7	98	2	JX0083	ferredoxin [2Fe-2S	628	5	3.7	109	2	S25038	Ig heavy chain V r
556	5	3.7	98	2	S22432	hydroxymethylgluta	629	5	3.7	109	2	B64145	hypothetical prote
557	5	3.7	98	2	T42909	hypothetical prote	630	5	3.7	109	2	D71209	probable ribosomal
558	5	3.7	98	2	T12887	hypothetical prote	631	5	3.7	109	2	C56413	major allergen Fel
559	5	3.7	99	2	B44866	hypothetical prote	632	5	3.7	109	2	B64492	hypothetical prote
560	5	3.7	99	2	A44866	major mezozote su	633	5	3.7	109	2	E84202	ferredoxin [import
561	5	3.7	99	2	A42196	preprotein translo	634	5	3.7	110	1	R5BYA1	acidic ribosomal p
562	5	3.7	99	2	D75129	hypothetical prote	635	5	3.7	110	2	A45938	transcription fact
563	5	3.7	99	2	T21919	hypothetical prote	636	5	3.7	110	2	S57412	Ig lambda chain V-
564	5	3.7	100	2	G84990	50S ribosomal prot	637	5	3.7	110	2	A32189	Ig heavy chain V r
565	5	3.7	100	2	P00818	expressed sequence	638	5	3.7	110	2	P00494	Ig heavy chain V r
566	5	3.7	100	2	C44866	major mezozote su	639	5	3.7	110	2	F69760	hypothetical prote
567	5	3.7	100	2	H85952	hypothetical prote	640	5	3.7	110	2	C71160	hypothetical prote
568	5	3.7	100	2	H65080	hypothetical prote	641	5	3.7	110	2	S74013	hypothetical prote
569	5	3.7	101	1	I64075	urease (EC 3.5.1.5	642	5	3.7	111	1	MEMS76	Ig heavy chain V-I
570	5	3.7	101	2	S26311	Ig heavy chain V r	643	5	3.7	111	2	S25052	Ig heavy chain V r
571	5	3.7	101	2	G70770	hypothetical prote	644	5	3.7	111	2	S25032	Ig heavy chain V r
572	5	3.7	101	2	S29093	SM3 protein - yea	645	5	3.7	111	2	S25047	Ig heavy chain V r
573	5	3.7	101	2	G96664	hypothetical prote	646	5	3.7	111	2	S25055	Ig heavy chain V r
574	5	3.7	101	2	E69442	conserved hypotet	647	5	3.7	111	2	S25048	Ig heavy chain V r
575	5	3.7	102	1	CCDV3D	cytochrome c3 - De	648	5	3.7	111	2	S25054	Ig heavy chain V r
576	5	3.7	102	2	S54825	glutaredoxin - cas	649	5	3.7	111	2	S25031	Ig heavy chain V r
577	5	3.7	102	2	PH1490	Ig heavy chain V r	650	5	3.7	111	2	S25051	Ig heavy chain V r
578	5	3.7	102	2	PH1491	Ig heavy chain V r	651	5	3.7	111	2	S25040	Ig heavy chain V r
579	5	3.7	102	2	F75077	hypothetical prote	652	5	3.7	111	2	S25033	Ig heavy chain V r
580	5	3.7	103	2	S28823	transcription fact	653	5	3.7	111	2	S25030	Ig heavy chain V r
581	5	3.7	104	1	CCCZ	cytochrome c - chi	654	5	3.7	111	2	S25045	Ig heavy chain V r
582	5	3.7	104	1	CCMOR	cytochrome c - rhe	655	5	3.7	111	2	S25034	Ig heavy chain V r
583	5	3.7	104	1	CCMKP	cytochrome c - spi	656	5	3.7	111	2	S25024	Ig heavy chain V r
584	5	3.7	104	2	G86739	50S ribosomal prot	657	5	3.7	111	2	G70772	hypothetical prote
585	5	3.7	104	2	S72391	hypothetical prote	658	5	3.7	112	1	R5FE2E	acidic ribosomal p
586	5	3.7	104	2	B71180	hypothetical prote	659	5	3.7	112	2	S09957	Ig heavy chain V-D
587	5	3.7	104	2	T13628	hypothetical prote	660	5	3.7	112	2	T38889	probable acyl carr
588	5	3.7	104	2	T13323	hypothetical prote	661	5	3.7	112	2	H69447	hypothetical prote
589	5	3.7	104	2	T17947	hypothetical prote	662	5	3.7	112	2	S58139	gene 9 protein - p
590	5	3.7	104	2	T36976	probable transposa	663	5	3.7	112	2	B48941	pediocin productio
591	5	3.7	105	1	CCHU	cytochrome c [vali	664	5	3.7	112	2	B86441	hypothetical prote
592	5	3.7	105	2	H69949	transcription regu	665	5	3.7	112	2	H70878	probable regulator
593	5	3.7	105	2	JQ2140	hypothetical 12.2K	666	5	3.7	112	2	B36259	Ig heavy chain V r
594	5	3.7	105	2	B47119	spore coat protein	667	5	3.7	113	2	S25041	Ig heavy chain V r
595	5	3.7	105	2	C81394	hypothetical prote	668	5	3.7	113	2	S25044	Ig heavy chain V r
596	5	3.7	105	2	H81293	hypothetical prote	669	5	3.7	113	2	D71205	hypothetical prote
597	5	3.7	105	2	E69326	hypothetical prote	670	5	3.7	113	2	I40399	flagellar protein
598	5	3.7	106	2	C69772	thioredoxin homolo	671	5	3.7	113	2	T19198	hypothetical prote
599	5	3.7	106	2	J50484	ribonuclease FI (E	672	5	3.7	114	2	T01262	Ig heavy chain V r
600	5	3.7	106	2	S13027	cysteine proteinas	673	5	3.7	114	2	S46377	Ig kappa chain V-J
601	5	3.7	106	2	S59639	Ig heavy chain V r	674	5	3.7	114	2	S11105	Ig heavy chain V r
602	5	3.7	106	2	S26464	Ig heavy chain V r	675	5	3.7	114	2	T08325	Ig heavy chain V r
603	5	3.7	106	2	S25036	Ig heavy chain V r	676	5	3.7	114	2	JC6077	hypothetical prote
604	5	3.7	106	2	H82094	bola protein WC229	677	5	3.7	114	2	T20095	p13 protein - Leuc
605	5	3.7	106	2	F71075	hypothetical prote	678	5	3.7	114	4	A47271	nitrophenyl phosph
606	5	3.7	106	2	H84333	hypothetical prote	679	5	3.7	115	2	S19965	Ig heavy chain V r
607	5	3.7	106	2	B64001	hypothetical prote	680	5	3.7	115	2	S19968	Ig heavy chain V r
608	5	3.7	106	2	S45762	hypothetical prote	681	5	3.7	115	2	S38714	Ig heavy chain V r
609	5	3.7	106	2	S32032	Sp17 protein precu	682	5	3.7	115	2	A54378	Ig heavy chain V r
610	5	3.7	107	2	A27646	Ig heavy chain V r	683	5	3.7	115	2	A56700	Ig heavy chain (an
611	5	3.7	107	2	PI0240	Ig heavy chain V r	684	5	3.7	115	2	S03482	Ig heavy chain V-D
612	5	3.7	107	2	PI0243	Ig heavy chain V r	685	5	3.7	115	2	S77491	ribosomal protein
613	5	3.7	107	2	PI0241	Ig heavy chain V r	686	5	3.7	115	2	A83776	hypothetical prote

687	5	3.7	115	2	S53390	probable membrane	760	5	3.7	121	2	S28738	hypothetical prote
688	5	3.7	116	2	S24289	Ig gamma chain V r	761	5	3.7	121	2	B86759	hypothetical prote
689	5	3.7	116	2	H29380	Ig gamma chain pre	762	5	3.7	121	2	T41406	very hypothetical
690	5	3.7	116	2	B28814	Ig heavy chain V r	763	5	3.7	121	2	T08717	hypothetical prote
691	5	3.7	116	2	S53751	antibody Fab Jel I	764	5	3.7	122	1	AVMSX2	Ig heavy chain V r
692	5	3.7	116	2	S44484	Ig heavy chain V r	765	5	3.7	122	2	PH1537	Ig H chain V regio
693	5	3.7	116	2	B75313	ribosomal protein	766	5	3.7	122	2	A45049	Ig heavy chain V r
694	5	3.7	116	2	C83951	ribosome-binding f	767	5	3.7	122	2	A27097	histone H2B.1 - Te
695	5	3.7	116	2	G71229	hypothetical prote	768	5	3.7	122	2	B27097	histone H2B.2 - Te
696	5	3.7	117	1	H71166	hypothetical prote	769	5	3.7	122	2	E86887	hypothetical prote
697	5	3.7	117	2	T13708	thiredoxin reduct	770	5	3.7	122	2	C64965	hypothetical prote
698	5	3.7	117	2	A45387	developmental regu	771	5	3.7	122	2	D71081	hypothetical prote
699	5	3.7	117	2	S32187	Ig heavy chain V r	772	5	3.7	122	2	B85825	hypothetical prote
700	5	3.7	117	2	S32190	Ig heavy chain V r	773	5	3.7	123	1	RSRT35	probable structura
701	5	3.7	117	2	PH1556	Ig H chain V regio	774	5	3.7	123	2	S60067	Ig heavy chain V r
702	5	3.7	117	2	S01822	Ig heavy chain V-D	775	5	3.7	123	2	G01477	ribosomal protein
703	5	3.7	117	2	S03305	Ig heavy chain V r	776	5	3.7	123	2	T43824	hypothetical prote
704	5	3.7	117	2	S25176	Ig heavy chain V r	777	5	3.7	123	2	S44555	probable membrane
705	5	3.7	117	2	S17586	Ig heavy chain V r	778	5	3.7	123	2	S56967	hypothetical prote
706	5	3.7	117	2	C81247	arsenate reductase	779	5	3.7	123	2	S27937	hypothetical prote
707	5	3.7	117	2	D64877	ycjd protein - Esc	780	5	3.7	123	2	T28245	hypothetical prote
708	5	3.7	117	2	S66720	probable membrane	781	5	3.7	123	2	T49377	hypothetical prote
709	5	3.7	117	2	B70572	hypothetical prote	782	5	3.7	123	2	S64248	hypothetical prote
710	5	3.7	117	2	T19098	hypothetical prote	783	5	3.7	123	2	D84322	hypothetical prote
711	5	3.7	117	2	T20888	hypothetical prote	784	5	3.7	123	2	T34403	hypothetical prote
712	5	3.7	118	2	B54025	transcriptlon fact	785	5	3.7	123	2	T32704	hypothetical prote
713	5	3.7	118	2	PH1529	Ig H chain V regio	786	5	3.7	124	2	PT0388	Ig heavy chain V r
714	5	3.7	118	2	PL0084	Ig heavy chain V r	787	5	3.7	124	2	PT0389	Ig heavy chain V r
715	5	3.7	118	2	S38565	Ig heavy chain V r	788	5	3.7	124	2	S20545	phosphate acceptor
716	5	3.7	118	2	S38717	Ig heavy chain V r	789	5	3.7	124	2	G84953	mutator mutR prote
717	5	3.7	118	2	PH0097	Ig heavy chain V r	790	5	3.7	124	2	G69821	hypothetical prote
718	5	3.7	118	2	E30540	Ig heavy chain V r	791	5	3.7	124	2	T37750	very hypothetical
719	5	3.7	118	2	PQ0266	Ig heavy chain V r	792	5	3.7	124	2	F42520	A3JR protein - vac
720	5	3.7	118	2	E69696	ribosomal protein	793	5	3.7	124	2	T45688	hypothetical prote
721	5	3.7	118	2	JC5915	hypothetical 13.3k	794	5	3.7	125	2	A31493	Ig light chain pre
722	5	3.7	118	2	T26728	hypothetical prote	795	5	3.7	125	2	T37417	probable 14.4K pro
723	5	3.7	118	2	C84859	hypothetical prote	796	5	3.7	125	2	D83548	hypothetical prote
724	5	3.7	118	2	C65078	hypothetical prote	797	5	3.7	125	2	T05609	hypothetical prote
725	5	3.7	119	2	H45722	anti-glycoprotein	798	5	3.7	126	2	S16280	Ig heavy chain (38
726	5	3.7	119	2	PH1541	Ig H chain V regio	799	5	3.7	126	2	G69132	ribosomal protein
727	5	3.7	119	2	F30502	Ig heavy chain V r	800	5	3.7	127	2	B72732	hypothetical prote
728	5	3.7	119	2	PH1531	Ig H chain V regio	801	5	3.7	127	2	T17431	hypothetical prote
729	5	3.7	119	2	PH1533	Ig H chain V regio	802	5	3.7	128	2	A37267	Ig heavy chain V r
730	5	3.7	119	2	B34353	anti-peptide Fab	803	5	3.7	128	2	C37267	Ig heavy chain V r
731	5	3.7	119	2	S03077	Ig heavy chain V r	804	5	3.7	128	2	I37267	Ig heavy chain V r
732	5	3.7	119	2	PH0098	Ig heavy chain V r	805	5	3.7	128	2	T15101	hypothetical prote
733	5	3.7	119	2	A24672	Ig heavy chain pre	806	5	3.7	129	2	P00062	T-cell receptor be
734	5	3.7	119	2	T16950	hypothetical prote	807	5	3.7	129	2	G69313	SSu ribosomal prot
735	5	3.7	119	2	T04243	acidic ribosomal p	808	5	3.7	129	2	JU0076	hypothetical 15K p
736	5	3.7	120	2	B42848	L6 mab heavy chain	809	5	3.7	129	2	T28490	hypothetical prote
737	5	3.7	120	2	A44371	Ig heavy chain V r	810	5	3.7	129	2	B72157	C1L protein - var
738	5	3.7	120	2	B2769	Ig heavy chain V r	811	5	3.7	129	2	E36842	EllL protein - var
739	5	3.7	120	2	B49715	Ig heavy chain V r	812	5	3.7	129	2	F42509	EllL 14.9K protein
740	5	3.7	120	2	PH1534	Ig H chain V regio	813	5	3.7	129	2	A44828	alkaline phosphata
741	5	3.7	120	2	S25175	Ig heavy chain V r	814	5	3.7	129	2	A81916	probable piliN NMA
742	5	3.7	120	2	G28195	Ig heavy chain V r	815	5	3.7	129	2	A81187	type IV piliN prot
743	5	3.7	120	2	F28195	Ig heavy chain V r	816	5	3.7	129	2	H71014	hypothetical prote
744	5	3.7	120	2	A61301	histone H2B - Tetr	817	5	3.7	130	1	RWMS20	T-cell receptor be
745	5	3.7	120	2	T31000	cysteine-rich prot	818	5	3.7	130	2	S73368	hypothetical prote
746	5	3.7	120	2	H86847	hypothetical prote	819	5	3.7	130	2	B86832	hypothetical prote
747	5	3.7	120	2	F84577	hypothetical prote	820	5	3.7	130	2	G96802	probable glutaredo
748	5	3.7	121	1	GVMS11	Ig heavy chain V r	821	5	3.7	130	2	S73018	hypothetical prote
749	5	3.7	121	1	S37200	Ig heavy chain V r	822	5	3.7	131	2	P00064	T-cell receptor be
750	5	3.7	121	2	F37266	Ig heavy chain V r	823	5	3.7	131	2	T40712	40S ribosomal prot
751	5	3.7	121	2	H37266	Ig heavy chain V r	824	5	3.7	131	2	T27224	hypothetical prote
752	5	3.7	121	2	PT0390	Ig heavy chain V r	825	5	3.7	131	2	J01767	SalIR protein - v
753	5	3.7	121	2	PT0391	Ig heavy chain V r	826	5	3.7	131	2	G66959	hypothetical prote
754	5	3.7	121	2	PL0281	Ig heavy chain V r	827	5	3.7	131	2	F75117	hypothetical prote
755	5	3.7	121	2	B34871	Ig heavy chain V r	828	5	3.7	132	1	HAFGH	hemoglobin alpha-t
756	5	3.7	121	2	A30551	Ig heavy chain V r	829	5	3.7	132	2	T50416	40S ribosomal prot
757	5	3.7	121	2	B69380	chemotaxis respons	830	5	3.7	132	2	S15966	hypothetical prote
758	5	3.7	121	2	A71132	hypothetical prote	831	5	3.7	132	2	F75108	hypothetical prote
759	5	3.7	121	2	T36836	hypothetical prote	832	5	3.7	132	2	G72351	conserved hypotnet



833	5	3.7	132	2	T03397	hypothetical prote	906	5	3.7	141	2	T09581	probable high mobi
834	5	3.7	133	2	PC1155	Ig heavy chain pre	907	5	3.7	142	2	C32394	succinate dehydrog
835	5	3.7	133	2	B81449	hypothetical prote	908	5	3.7	142	2	S71361	actin-binding prot
836	5	3.7	133	2	S69803	hypothetical prote	909	5	3.7	142	2	G64952	hypothetical 17.1
837	5	3.7	134	2	S73181	H+-transporting AT	910	5	3.7	142	2	H85802	probable regulator
838	5	3.7	134	2	E57233	complexin II - hum	911	5	3.7	142	2	D75134	hypothetical prote
839	5	3.7	134	2	D57233	complexin II - mou	912	5	3.7	142	2	D72040	conserved hypotet
840	5	3.7	134	2	JC4226	synaphin - bovine	913	5	3.7	142	2	C86583	Ct635 hypothetical
841	5	3.7	134	2	S66294	921-S protein - mo	914	5	3.7	142	2	E84369	histidine triad pr
842	5	3.7	134	2	A57233	complexin I - rat	915	5	3.7	142	2	S46445	YMF46 protein - Ac
843	5	3.7	134	2	C57233	complexin II - rat	916	5	3.7	142	2	T44136	hypothetical prote
844	5	3.7	134	2	A82455	hypothetical prote	917	5	3.7	143	1	RGMX15	ribosomal protein
845	5	3.7	134	2	G82477	hypothetical prote	918	5	3.7	143	2	B22394	succinate dehydrog
846	5	3.7	134	2	E70251	hypothetical prote	919	5	3.7	143	2	D69083	hypothetical prote
847	5	3.7	134	2	E75582	hypothetical prote	920	5	3.7	143	2	F86736	general stress pro
848	5	3.7	134	2	T12219	glutaredoxin I - c	921	5	3.7	143	2	T31273	hypothetical prote
849	5	3.7	135	1	PWLVE	H+-transporting AT	922	5	3.7	144	2	S23658	superoxide dismuta
850	5	3.7	135	2	PL0100	Ig heavy chain pre	923	5	3.7	144	2	B30502	Ig heavy chain V r
851	5	3.7	135	2	S31913	Ig gamma-2A chain	924	5	3.7	144	2	T43374	ribosomal protein
852	5	3.7	135	2	I56326	fatty acid binding	925	5	3.7	144	2	T41959	neutrophil-activat
853	5	3.7	136	1	HVMSB1	Ig heavy chain pre	926	5	3.7	144	2	C64550	neutrophil activat
854	5	3.7	136	1	R5EC16	ribosomal protein	927	5	3.7	144	2	B72214	hypothetical prote
855	5	3.7	136	2	PL0077	Ig heavy chain pre	928	5	3.7	144	2	C70455	hypothetical prote
856	5	3.7	136	2	PL0208	Ig heavy chain pre	929	5	3.7	144	2	A69399	hypothetical prote
857	5	3.7	136	2	B47159	Ig heavy chain V r	930	5	3.7	144	2	C83974	hypothetical prote
858	5	3.7	136	2	S13791	Ig heavy chain V r	931	5	3.7	144	2	E86618	hypothetical prote
859	5	3.7	136	2	T30159	hypothetical prote	932	5	3.7	144	2	A84168	hypothetical prote
860	5	3.7	136	2	C64093	ribosomal protein	933	5	3.7	144	2	T08666	ribosomal protein
861	5	3.7	136	2	F85996	50S ribosomal subu	934	5	3.7	145	1	R5RT26	transcription regu
862	5	3.7	136	2	G82058	ribosomal protein	935	5	3.7	145	1	G69867	superoxide dismuta
863	5	3.7	136	2	S73520	MG441 homolog E09-	936	5	3.7	145	2	S54794	Ig heavy chain pre
864	5	3.7	136	2	T22240	hypothetical prote	937	5	3.7	145	2	S03844	ribosomal protein
865	5	3.7	136	2	T22797	hypothetical prote	938	5	3.7	145	2	S33713	ribosomal protein
866	5	3.7	137	1	G2MS43	Ig heavy chain pre	939	5	3.7	145	2	S48864	hypothetical prote
867	5	3.7	137	1	AVMS35	Ig heavy chain pre	940	5	3.7	145	2	T33961	hypothetical prote
868	5	3.7	137	2	H32513	Ig heavy chain pre	941	5	3.7	145	2	T13550	hypothetical prote
869	5	3.7	137	2	B64208	hypothetical prote	942	5	3.7	145	2	H84174	hypothetical prote
870	5	3.7	137	2	B70123	conserved hypotet	943	5	3.7	145	2	B72724	hypothetical prote
871	5	3.7	137	2	G71551	hypothetical prote	944	5	3.7	145	2	S37043	Salt protein precu
872	5	3.7	137	2	T22872	hypothetical prote	945	5	3.7	145	2	S19929	microtubule-assoc
873	5	3.7	137	2	S54618	hypothetical prote	946	5	3.7	146	1	S24311	flavodoxin - Desul
874	5	3.7	138	2	G81231	50S ribosomal prot	947	5	3.7	146	1	R3MX7	ribosomal protein
875	5	3.7	138	2	S24103	envelope protein -	948	5	3.7	146	2	B64246	ribosomal protein
876	5	3.7	138	2	T02663	ascorbic acid- and	949	5	3.7	146	2	H72167	A35R protein - var
877	5	3.7	138	2	G82116	flagellar motor sw	950	5	3.7	146	2	S73456	DNA-directed RNA p
878	5	3.7	138	2	S36115	interferon - Japan	951	5	3.7	146	2	G70020	hypothetical prote
879	5	3.7	139	1	MHMS18	Ig heavy chain pre	952	5	3.7	146	4	S33905	Ig heavy chain pre
880	5	3.7	139	2	A61089	beta-crystallin B3	953	5	3.7	147	2	F84153	ribosomal protein
881	5	3.7	139	2	H64302	conserved hypotet	954	5	3.7	147	2	D75132	probable hydrogena
882	5	3.7	139	2	F82863	conserved hypotet	955	5	3.7	147	2	F71089	probable hydrogena
883	5	3.7	139	2	C83807	hypothetical prote	956	5	3.7	147	2	A69933	hypothetical prote
884	5	3.7	140	2	PH1482	Ig heavy chain V r	957	5	3.7	147	2	G70795	hypothetical prote
885	5	3.7	140	2	PH1483	Ig heavy chain V r	958	5	3.7	147	2	F83801	chordimate mutase
886	5	3.7	140	2	PH1484	Ig heavy chain V r	959	5	3.7	147	2	H82172	heat shock protein
887	5	3.7	140	2	PH1488	Ig heavy chain V r	960	5	3.7	147	2	B69780	conserved hypotet
888	5	3.7	140	2	PH1489	Ig heavy chain V r	961	5	3.7	148	1	T12787	probable protein-d
889	5	3.7	140	2	PH1498	Ig heavy chain V r	962	5	3.7	148	2	A57569	tauinstar protein -
890	5	3.7	140	2	S06816	Ig heavy chain pre	963	5	3.7	148	2	C82544	limbriar protein x
891	5	3.7	140	2	A36194	Ig heavy chain V r	964	5	3.7	148	2	T33812	hypothetical prote
892	5	3.7	140	2	B69496	small heat shock p	965	5	3.7	148	2	A72785	hypothetical prote
893	5	3.7	140	2	S31628	MA28L protein - my	966	5	3.7	148	2	C75212	hypothetical prote
894	5	3.7	140	2	H69405	hypothetical prote	967	5	3.7	148	2	T22508	hypothetical prote
895	5	3.7	140	2	T28576	6R protein - vario	968	5	3.7	149	2	H70441	ribosomal protein
896	5	3.7	140	2	H36851	RNA-binding ribonu	969	5	3.7	149	2	S36992	transposase (clone
897	5	3.7	140	2	E70632	hypothetical prote	970	5	3.7	149	2	B37169	TUD4 protein precu
898	5	3.7	141	2	JL0076	endonuclease V-lik	971	5	3.7	149	2	S39556	high mobility grou
899	5	3.7	141	2	A39276	Ig heavy chain pre	972	5	3.7	149	2	G64414	hypothetical prote
900	5	3.7	141	2	S08118	histone H2A.VD - f	973	5	3.7	149	2	A54503	51k merozoit surf
901	5	3.7	141	2	S75127	hypothetical prote	974	5	3.7	150	2	H86646	deoxyuridine 5'-tr
902	5	3.7	141	2	S74873	hypothetical prote	975	5	3.7	150	2	H86194	hypothetical prote
903	5	3.7	141	2	S77513	hypothetical prote	976	5	3.7	150	2	B26368	protein-serine kin
904	5	3.7	141	2	E72781	hypothetical prote	977	5	3.7	150	2	F59105	hypothetical prote
905	5	3.7	141	2			978	5	3.7	151	2	S77475	ribosomal protein

979 5 3.7 151 2 A69227  
 980 5 3.7 151 2 A35735  
 981 5 3.7 151 2 C68352  
 982 5 3.7 152 1 C69202  
 983 5 3.7 152 2 B64369  
 984 5 3.7 152 2 B64485  
 985 5 3.7 152 2 T02703  
 986 5 3.7 152 2 A70850  
 987 5 3.7 153 2 A64333  
 988 5 3.7 153 2 S54841  
 989 5 3.7 153 2 B60891  
 990 5 3.7 153 2 T04614  
 991 5 3.7 153 2 E69403  
 992 5 3.7 153 2 D83012  
 993 5 3.7 154 2 C81907  
 994 5 3.7 154 2 F81110  
 995 5 3.7 154 2 S25994  
 996 5 3.7 154 2 A05187  
 997 5 3.7 154 2 C72503  
 998 5 3.7 154 2 T47825  
 999 5 3.7 154 2 S39729  
 1000 5 3.7 154 2 T03503

## ALIGNMENTS

hypothetical prote  
 regulatory protein  
 protein T26F17.12  
 conserved hypotet  
 hypothetical prote  
 hypothetical prote  
 hypothetical prote  
 hypothetical prote  
 formate hydrogenly  
 peptidylprolyl iso  
 plin - Dichelobac  
 hypothetical prote  
 hypothetical prote  
 hypothetical prote  
 probable bacteriof  
 bacterioferritin A  
 hypothetical prote  
 hypothetical prote  
 Skpl-like protein  
 ywea protein - Bac  
 conserved hypotet

RESULT 1  
 PRSAK  
 staphylokinase - phage S phl-C  
 C:Species: phage S phl-C  
 A:Note: host Staphylococcus aureus  
 C:Date: 28-Aug-1985 #sequence\_revision 28-Aug-1985 #text\_change 28-May-1999  
 C:Accession: A00995  
 R:Sako, T.; Tsuchida, N.  
 Nucleic Acids Res. 11, 7679-7693, 1983  
 A:Title: Nucleotide sequence of the staphylokinase gene from Staphylococcus aureus.  
 A:Reference number: A00995; MUID:84069795  
 A:Accession: A00995  
 A:Molecule type: DNA  
 A:Residues: 1-163 <SAK>  
 A:Cross-references: GB:X00127; NID:947425; PIDN:CAA24957.1; PID:G758303  
 C:Comment: Although it has no intrinsic proteolytic activity, this secreted protein is a  
 ructure of plasminogen is thought to expose the plasmin active site.  
 C:Comment: The designation of staphylokinase as synonymous with Staphylococcal aureus ne  
 C:Genetics:  
 A:Gene: sak  
 A:Superfamily: phage S phl-C staphylokinase  
 C:Keywords: plasminogen activator

Query Match 75.0%; Score 102; DB 1; Length 163;  
 Best Local Similarity 100.0%; Pred. No. 7.1e-99;  
 Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 35 KGNELLSPIYVEPIKPGTTLTKKEIEYVEMALDATAKERRVVELDPSAKIEVYTDK 94  
 |||||||  
 DB 62 KGNELLSPIYVEPIKPGTTLTKKEIEYVEMALDATAKERRVVELDPSAKIEVYTDK 121  
 |||||||

OY 95 NKKKEETKSPITTEKGFVVPDLSEHKINPGFNLTIKVIEKK 136  
 |||||||  
 DB 122 NKKKEETKSPITTEKGFVVPDLSEHKINPGFNLTIKVIEKK 163  
 |||||||

RESULT 2  
 S02330  
 staphylokinase - phage P42D  
 C:Species: phage P42D  
 A:Note: host Staphylococcus aureus  
 C:Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 26-Aug-1999  
 C:Accession: S02330; S45654  
 R:Behnke, D.; Gerlach, D.  
 Mol. Gen. Genet. 210, 528-534, 1987

A:Title: Cloning and expression in Escherichia coli, Bacillus subtilis, and Streptoco  
 A:Reference number: S02330; MUID:88121731  
 A:Accession: S02330  
 A:Molecule type: DNA  
 A:Residues: 1-163 <BEH>  
 A:Cross-references: EMBL:X06603; NID:946676; PIDN:CAA29822.1; PID:G758275  
 R:Gase, A.; Birch-Hirschfeld, E.; Guehrs, K.H.; Hartmann, M.; Vetterman, S.; Damaschu  
 Eur. J. Biochem. 223, 303-308, 1994  
 A:Title: The thermostability of natural variants of bacterial plasminogen-activator S  
 A:Reference number: S45654; MUID:94307274  
 A:Accession: S45654  
 A:Molecule type: protein  
 A:Residues: 28-163 <GAS>  
 C:Genetics:  
 A:Gene: sak  
 C:Superfamily: phage S phl-C staphylokinase

Query Match 68.4%; Score 93; DB 2; Length 163;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-89;  
 Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 44 YVEPIKPGTTLTKKEIEYVEMALDATAKERRVVELDPSAKIEVYTDK 103  
 |||||||  
 DB 71 YVEPIKPGTTLTKKEIEYVEMALDATAKERRVVELDPSAKIEVYTDK 130  
 |||||||

OY 104 FPIITEKGFVVPDLSEHKINPGFNLTIKVIEKK 136  
 |||||||  
 DB 131 FPIITEKGFVVPDLSEHKINPGFNLTIKVIEKK 163  
 |||||||

RESULT 3  
 A72218  
 gufa protein - Thermotoga maritima (strain MSB8)  
 C:Species: Thermotoga maritima  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
 C:Accession: A72218  
 R:Nejlon, K.E.; Clayton, R.A.; Gill, S.R.; Gwyn, M.L.; Dodson, R.J.; Haft, D.H.; Hic  
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,  
 C.M.  
 Nature 399, 323-329, 1999  
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome  
 A:Reference number: A72200; MUID:99287316  
 A:Accession: A72218  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-245 <ARN>  
 A:Cross-references: GB:AE001812; GB:AE000512; NID:94982302; PIDN:AD36803.1; PID:9498  
 A:Experimental source: strain MSB8  
 C:Genetics:  
 A:Gene: TM1738  
 C:Superfamily: gufa protein

Query Match 5.9%; Score 8; DB 2; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 1.5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 34 SKGNELLS 41  
 |||||||  
 DB 219 SKGNELLS 226  
 |||||||

RESULT 4  
 C69044  
 hypothetical protein MTH1333 - Methanobacterium thermoautotrophicum (strain Delta H)  
 C:Species: Methanobacterium thermoautotrophicum  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
 C:Accession: C69044  
 R:Smith, D.R.; Doucelte-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T  
 Qiu, D.; Spadafora, R.; Vicalre, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani,  
 Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
 J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funcn  
A:Reference number: A69000; MUID:98037514  
A:Accession: C69044  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-56 <MTH>  
A:Cross-references: GB:AE000897; GB:AE000666; NID:92622439; PIDN:AB85811.1; PID:9262244  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH1333

Query Match 5.1%; Score 7; DB 2; Length 56;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 53 TTTTKEK 59  
|||||||  
Db 16 TTTTKEK 22

RESULT 5  
I46953  
CGM-gated retinal photoreceptor channel - rabbit (fragment)  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 04-Sep-1997 #sequence\_revision 07-Nov-1997 #text\_change 13-Aug-1999  
C:Accession: I46953  
R:Hundal, S.P.; DiFrancesco, D.; Mangoni, M.; Brammar, W.J.; Conley, E.C.  
Biochem. Soc. Trans. 21, 119S, 1993.  
A:Title: An isoform of the CGM-gated retinal photoreceptor channel gene expressed in th  
A:Reference number: I46953; MUID:93365767  
A:Accession: I46953  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-90 <HUV>  
A:Cross-references: GB:565218; NID:9410551; PIDN:AB827924.1; PID:9410552  
C:Superfamily: cyclic nucleotide-gated channel; cAMP receptor protein cyclic nucleotide-

Query Match 5.1%; Score 7; DB 2; Length 90;  
Best Local Similarity 100.0%; Pred. No. 6.9;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 96 KKKERTK 102  
|||||||  
Db 6 KKKERTK 12

RESULT 6  
PQ0629  
coat protein - tulip top-breaking virus (fragment)  
C:Species: tulip top-breaking virus  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 17-Nov-2000  
C:Accession: PQ0629  
R:DeKker, E.L.; Derks, A.F.L.M.; Asjes, C.J.; Lemmers, M.E.C.; Bol, J.F.; Langeveld, S.A.  
J. Gen. Virol. 74, 881-887, 1993  
A:Title: Characterization of polylviruses from tulip and lily which cause flower-breaking  
A:Reference number: PQ0628; MUID:93260402  
A:Accession: PQ0629  
A:Molecule type: mRNA  
A:Residues: 1-92 <DEK>  
A:Cross-references: GB:S60806; NID:9385510; PIDN:AB826636.1; PID:9385511  
C:Superfamily: tobacco etch virus genome polypeptide  
C:Keywords: coat protein

Query Match 5.1%; Score 7; DB 2; Length 92;  
Best Local Similarity 100.0%; Pred. No. 7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 VEEPIKP 51  
|||||||  
Db 18 VEEPIKP 24

RESULT 7  
C84026  
ribosomal protein L21 (BU20) rplU [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 08-Dec-2000  
C:Accession: C84026  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a  
A:Reference number: A83650; MUID:20263314  
A:Accession: C84026  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-102 <STO>  
A:Cross-references: GB:AP001517; GB:BA000004; NID:910175500; PIDN:BA806730.1; GSPDB:G  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: rplU  
C:Superfamily: Escherichia coli ribosomal protein L21

Query Match 5.1%; Score 7; DB 2; Length 102;  
Best Local Similarity 100.0%; Pred. No. 7.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 129 TKVIEK 135  
|||||||  
Db 93 TKVIEK 99

RESULT 8  
D83704  
transposase (14) BH0436 [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 31-Dec-2000  
C:Accession: D83704  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a  
A:Reference number: A83650; MUID:20263314  
A:Accession: D83704  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-138 <STO>  
A:Cross-references: GB:AP001508; GB:BA000004; NID:910172890; PIDN:BA804155.1; GSPDB:G  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH0436

Query Match 5.1%; Score 7; DB 2; Length 138;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 DSKGNEL 39  
|||||||  
Db 80 DSKGNEL 86

RESULT 9  
E71501  
hypothetical protein CT548 - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
C:Species: Chlamydia trachomatis  
C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 08-Oct-1999  
C:Accession: E71501  
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche  
Science 282, 754-759, 1998  
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t  
A:Reference number: A71570; MUID:99000809  
A:Accession: E71501  
A:Status: preliminary

A:Molecule type: DNA  
 A:Residues: 1-194 <ARN>  
 A:Cross-references: GB:AE001325; GB:AE001273; NID:g33328980; PIDN:AAC68150.1; PID:g3332898  
 A:Experimental source: serotype D, strain UW-3/CX  
 C:Genetics:  
 A:Gene: Ctr548

Query Match 5.1%; Score 7; DB 2; Length 194;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 40 LSPHYE 46  
 Db 25 LSPHYE 31

## RESULT 10

tryptophan synthase alpha chain [imported] - Halobacterium sp. NRC-1  
 C:Species: Halobacterium sp. NRC-1  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: D84190  
 R:Ng, W.V.; Kennedy, S.P.; Mahapatra, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Lehtsamer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jadhav, K.H.; Alam, M.; Freitas, F.; Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li  
 A:Title: Genome sequence of Halobacterium species NRC-1.  
 A:Reference number: A84160; MUID:20504483  
 A:Accession: D84190  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-204 <STO>  
 A:Cross-references: GB:AE004437; NID:g10579936; PIDN:AG18888.1; GSPDB:GN00138  
 C:Genetics:  
 A:Gene: trpA

Query Match 5.1%; Score 7; DB 2; Length 204;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 110 GFVPPDL 116  
 Db 51 GFVPPDL 57

## RESULT 11

hypothetical protein Y6E2A.8 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T27336  
 R:Matthews, L.  
 submitted to the EMBL Data Library, January 1998  
 A:Reference number: Z20347  
 A:Accession: T27336  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-219 <MIL>  
 A:Cross-references: EMBL:AL021175; PIDN:CA15969.1; GSPDB:GN00023; CESP:Y6E2A.8  
 A:Experimental source: clone Y6E2A  
 C:Genetics:  
 A:Gene: CESP:Y6E2A.8  
 A:Map position: 5

Query Match 5.1%; Score 7; DB 2; Length 219;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 93 DKNKKE 99

Db 126 DKNKKE 132

## RESULT 12

tryptophan synthase (EC 4.2.1.20) alpha chain [validated] - Haloferax volcanii (strain A36044)  
 C:Species: Haloferax volcanii  
 C:Date: 25-Jan-1991 #sequence\_revision 25-Jan-1991 #text\_change 26-May-2000  
 C:Accession: A36044  
 R:Ham, W.L.; Cohen, A.; Tsoulfas, D.; Doolittle, W.F.; Proc. Natl. Acad. Sci. U.S.A. 87, 6614-6618, 1990  
 A:Title: Genes for tryptophan biosynthesis in the archaeobacterium Haloferax volcanii.  
 A:Reference number: A36044; MUID:90370836  
 A:Accession: A36044  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-217 <LAM>  
 A:Cross-references: GB:M6177; NID:g149041; PIDN:AAA72864.1; PID:g149044  
 C:Genetics:  
 A:Gene: trpA  
 C:Function:  
 A:Description: EC 4.2.1.20 [validated; MUID:90370836]  
 C:Superfamily: tryptophan synthase alpha chain; tryptophan synthase alpha chain homol  
 C:Keywords: carbon-oxygen lyase; hydro-lyase  
 F:12-240/Domain: tryptophan synthase alpha chain homology <TRPA>  
 F:43/Active site: Glu #status predicted

Query Match 5.1%; Score 7; DB 2; Length 277;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 110 GFVPPDL 116  
 Db 120 GFVPPDL 126

## RESULT 13

coat protein - turnip mosaic virus (strain UK1) (fragment)  
 C:Species: turnip mosaic virus, TURN  
 A:Variety: strain UK1  
 C:Date: 20-Feb-1995 #sequence\_revision 26-Apr-1996 #text\_change 17-Nov-2000  
 C:Accession: S21499  
 R:Lehmann, P.; Kozubek, E.; Ostrowska, K.; Walsh, J.; Greenland, A.  
 submitted to the EMBL Data Library, April 1992  
 A:Description: Nucleotide sequence of the coat protein gene of Turnip mosaic virus tr  
 A:Reference number: S21499  
 A:Accession: S21499  
 A:Molecule type: DNA  
 A:Residues: 1-288 <LEH>  
 A:Cross-references: EMBL:X65978  
 A:Experimental source: strain UK1  
 C:Superfamily: tobacco etch virus genome polypeptide  
 C:Keywords: coat protein

Query Match 5.1%; Score 7; DB 2; Length 288;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 VEPPIKP 51  
 Db 162 VEPPIKP 168

## RESULT 14

S51329  
 polyprotein - turnip mosaic virus (strain cgs) (fragment)  
 N:Contains: coat protein  
 C:Species: turnip mosaic virus, TURN  
 A:Variety: strain cgs

C;Date: 19-Mar-1997 #sequence\_revision 30-Jan-1998 #text\_change 17-Nov-2000  
 C;Accession: S51329  
 R;Choi, G.S.; Choi, J.K.  
 A;Submitted to the EMBL Data Library, January 1995  
 A;Description: Nucleotide sequence of coat protein gene of turnip mosaic virus (cgs strain)  
 A;Reference number: S51329  
 A;Accession: S51329  
 A;Molecule type: genomic RNA  
 A;Residues: 1-288 <CHO>  
 A;Cross-references: EMBL:X83968; NID:g1552715; PIDN:CAA58802.1; PID:g634105  
 A;Experimental source: strain cgs  
 C;Superfamily: tobacco etch virus genome polyprotein  
 C;Keywords: coat protein; polyprotein  
 F;1-288/Product: coat protein #status predicted <MAT>

Query Match 5.1%; Score 7; DB 2; Length 288;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 VEPPIKP 51  
 |||||  
 Db 162 VEPPIKP 168

RESULT 15  
 A83746  
 transposase (16) BH0769 [imported] - Bacillus halodurans (strain C-125)  
 C;Species: Bacillus halodurans  
 C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 31-Dec-2000  
 C;Accession: A83746  
 R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
 A;Reference number: A83650; MUID:20263314  
 A;Accession: A83746  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-311 <STO>  
 A;Cross-references: GB:AP001509; GB:BA000004; NID:g10173176; PIDN:BA04488.1; GSPDB:GN00  
 A;Experimental source: strain C-125  
 C;Genetics:  
 A;Gene: BH0769

Query Match 5.1%; Score 7; DB 2; Length 311;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 DSKGNEL 39  
 |||||  
 Db 262 DSKGNEL 268

Search completed: April 22, 2002, 10:46:10  
 Job time: 130 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 10:45:45 ; Search time 12.01 Seconds

(without alignments)  
415.189 Million cell updates/sec

Title: us-09-601-490-1

Perfect score: 136  
Sequence: 1 SSSFDKGGKKYKGDASTFEP.....SEHIKNPGFNLITRVIEKK 136

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 100059 seqs, 36664827 residues

Word size : 0

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	102	75.0	163	1	SAK_STRAU
2	93	68.4	163	1	SAK_BPP42
3	7	5.1	204	1	TRPA_HA1N1
4	7	5.1	231	1	CTRA_CAUCR
5	7	5.1	277	1	TRPA_HA1VO
6	7	5.1	334	1	TRXB_NEUCR
7	7	5.1	722	1	PBP1_YEAST
8	7	5.1	944	1	IF2_HELPY
9	7	5.1	1034	1	POL_HY2CA
10	7	5.1	1035	1	POL_HY2KR
11	7	5.1	1035	1	POL_HY2NR
12	7	5.1	1036	1	POL_HY2RO
13	7	5.1	1049	1	POL_HY2G1
14	7	5.1	1055	1	POL_HY2ST
15	7	5.1	1073	1	POL_HY2D1
16	7	5.1	1142	1	POL_HY2BE
17	7	5.1	3163	1	POLG_TUWVQ
18	7	5.1	3164	1	POLG_TUWVJ
19	6	4.4	63	1	YPTA_BACSU
20	6	4.4	85	1	YKK2_CAEEL
21	6	4.4	90	1	POLG_GSMV
22	6	4.4	128	1	GCSSH_ECOLI
23	6	4.4	137	1	RK19_NEPOL
24	6	4.4	138	1	RBS_GALISU
25	6	4.4	146	1	PABZ_PSEAE
26	6	4.4	153	1	HYCA_ECOLI
27	6	4.4	153	1	YORZ_LISMO
28	6	4.4	158	1	RR7_ODOST
29	6	4.4	160	1	RS7A_AQUAE
30	6	4.4	160	1	RS7B_AQUAE
31	6	4.4	175	1	HER1_CAEEL
32	6	4.4	183	1	VH96_ARCFU
33	6	4.4	193	1	SAP3_HUMAN
34	6	4.4	220	1	GTC1_RAT
35	6	4.4	223	1	GTA1_RABIT
36	6	4.4	227	1	FLPA_PYRAB
37	6	4.4	227	1	FLPA_PYRAB
38	6	4.4	230	1	UBL3_HUMAN
39	6	4.4	236	1	GRA7_TOXGO
40	6	4.4	237	1	MINC_BUCAT
41	6	4.4	240	1	LEC_BOWMI
42	6	4.4	260	1	NMA_HUMAN
43	6	4.4	269	1	CART_CHICK
44	6	4.4	270	1	IF2B_WHEAT
45	6	4.4	270	1	PP12_HUMAN
46	6	4.4	270	1	PP12_MOUSE
47	6	4.4	270	1	PP12_RAT
48	6	4.4	279	1	Y011_BORBU
49	6	4.4	281	1	MODD_HAELN
50	6	4.4	283	1	DMPD_PSESP
51	6	4.4	284	1	COAT_BPLIH
52	6	4.4	303	1	NODD_RHILY
53	6	4.4	313	1	NOD3_RHIME
54	6	4.4	320	1	FLA1_TREHY
55	6	4.4	322	1	NODD_RHILE
56	6	4.4	323	1	CD47_HUMAN
57	6	4.4	326	1	YACJ_RHISN
58	6	4.4	327	1	MOA1_SYNT3
59	6	4.4	330	1	REBI_SALTY
60	6	4.4	340	1	ASPG_FLAME
61	6	4.4	345	1	HEMA_INBRU
62	6	4.4	345	1	HEMA_INBYB
63	6	4.4	345	1	YDE9_SCHPO
64	6	4.4	347	1	HEMA_INBFI
65	6	4.4	347	1	HEMA_INBF2
66	6	4.4	347	1	HEMA_INBF3
67	6	4.4	347	1	HEMA_INBF6
68	6	4.4	347	1	HEMA_INBF9
69	6	4.4	347	1	HEMA_INBFI
70	6	4.4	347	1	HEMA_INBNA
71	6	4.4	353	1	HEMA_INBNA
72	6	4.4	353	1	PSBA_MESV1
73	6	4.4	361	1	MTD_MESCR
74	6	4.4	363	1	HEMA_INBGL
75	6	4.4	365	1	HEMA_INBSJ
76	6	4.4	365	1	HEMA_INBID
77	6	4.4	368	1	ILVE_MYCTU
78	6	4.4	377	1	MTC3_CHVNI
79	6	4.4	392	1	BCAM_HUMAN
80	6	4.4	392	1	GCST_CHICK
81	6	4.4	392	1	RPA2_SULAC
82	6	4.4	393	1	ILIS_CERAE
83	6	4.4	396	1	CSA_HUMAN
84	6	4.4	396	1	EFPU_MYCLE
85	6	4.4	397	1	EFPU_PLARO
86	6	4.4	397	1	RL3_DICDI
87	6	4.4	398	1	ILIS_HUMAN
88	6	4.4	407	1	Y173_AQUAE
89	6	4.4	412	1	P56_YEAST
90	6	4.4	416	1	YMG5_CAEEL
91	6	4.4	425	1	YGGP_ECOLI
92	6	4.4	428	1	PROB_YEAST
93	6	4.4	434	1	TAF7_TREPA
94	6	4.4	437	1	PAK_ECOLI
95	6	4.4	452	1	AAHP_HUMAN
96	6	4.4	457	1	GAL8_KLUTLA
97	6	4.4	462	1	CATC_MOUSE
98	6	4.4	469	1	YH89_METUA
99	6	4.4	480	1	HD11_XENLA
100	6	4.4	480	1	HD12_XENLA
101	6	4.4	480	1	HD41_CHICK
102	6	4.4	481	1	6PGD_DROME
103	6	4.4	481	1	6PGD_DROST
104	6	4.4	482	1	ALAA2_PANMI
105	6	4.4	487	1	YOM5_CAEEL
106	6	4.4	489	1	PABA_BACSU
P04904	rattus norv				
O08863	oryzolaqus				
O9V215	pyrococcus				
O57811	pyrococcus				
P15374	homo sapien				
O00933	toxoplasma				
P57412	buchnera ap				
P42088	bowringia m				
O13145	homo sapien				
P07090	gallus gall				
O24473	trititum ae				
P48739	homo sapien				
P33811	mus musculu				
P53812	rattus norv				
O51044	borrelia bu				
O57278	haemophilus				
P19076	pseudomonas				
O04754	lactococcus				
P04681	rhizobium l				
P23190	rhizobium m				
P32520	treponema h				
P16556	rhizobium l				
O08722	homo sapien				
P55392	rhizobium s				
O55369	synectocyst				
P26395	salmonella				
O47898	flavobacter				
P18877	influenza b				
P18880	influenza b				
O10442	schizosacch				
O07925	influenza b				
O07926	influenza b				
O07922	influenza b				
O07924	influenza b				
O07920	influenza b				
P18878	influenza b				
P18879	influenza b				
P12440	influenza b				
O9nuw0	mesostigma				
P93257	mesembryant				
P12442	influenza b				
P12443	influenza b				
P12441	influenza b				
O10399	mycobacteri				
P10835	chlorella v				
O15382	homo sapien				
P28337	gallus gall				
P11514	sulfolobus				
O29612	cercopithec				
O12126	homo sapien				
P30768	mycobacteri				
P72231	planoblastor				
P34113	dictyostell				
P27930	homo sapien				
O67720	aquifex aeo				
P25270	saccharomyc				
O20932	caenorhabdi				
P52048	escherichia				
P32264	saccharomyc				
P29723	treponema p				
P76085	escherichia				
O13685	homo sapien				
O06433	kluyveromyc				
P97821	mus musculu				
O58299	methanococ				
O91695	xenopus lae				
O42227	xenopus lae				
P55517	gallus gall				
P41572	drosophila				
P41573	drosophila				
P34106	pancum m11				
P30651	caenorhabdi				
P37967	baecillus su				

107	6	4.4	500	1	ANP1_YEAST	P32629 saccharomyc	180	6	4.4	950	1	CDAD_HUMAN	O9y513 homo sapien
108	6	4.4	526	1	CAP_YEAST	P17555 saccharomyc	181	6	4.4	950	1	CDAD_HUMAN	O9y518 homo sapien
109	6	4.4	529	1	YL27_CAEEL	P34470 caenorhabditi	182	6	4.4	950	1	CDAD_HUMAN	O9un73 homo sapien
110	6	4.4	534	1	CH60_GALISU	P28256 galdieria s	183	6	4.4	950	1	CDAD_HUMAN	O9y515 homo sapien
111	6	4.4	536	1	TMAS_HUMAN	O15131 homo sapien	184	6	4.4	950	1	CDAD_HUMAN	O9y510 homo sapien
112	6	4.4	536	1	INR2_SHEEP	O95207 ovis aries	185	6	4.4	950	1	CDAD_HUMAN	O9y515 homo sapien
113	6	4.4	533	1	YF61_METJA	O58956 methanococc	186	6	4.4	963	1	CDCI_HUMAN	O9h158 homo sapien
114	6	4.4	574	1	HEMA_INBMD	P03461 influenza b	187	6	4.4	966	1	PHSI_SOLTU	P04045 homo sapien
115	6	4.4	575	1	HEMA_INBBO	P10448 influenza b	188	6	4.4	1007	1	CDIC_HUMAN	O9y514 homo sapien
116	6	4.4	575	1	HEMA_INBHK	P03462 influenza b	189	6	4.4	1018	1	HMW1_MYCPN	O50365 mycoplasma
117	6	4.4	576	1	HEMA_INBMS	P09766 influenza b	190	6	4.4	1025	1	ADAI_YEAST	O50365 mycoplasma
118	6	4.4	578	1	HEMA_INBME	P09765 influenza b	191	6	4.4	1120	1	RPMO_SCHPO	O13993 schizosacch
119	6	4.4	578	1	HEMA_INBVI	P09767 influenza b	192	6	4.4	1132	1	TERP_HUMAN	O14746 homo sapien
120	6	4.4	583	1	HEMA_INBOR	P10757 influenza b	193	6	4.4	1211	1	DP3A_HELPJ	O9z4f9 helicobacte
121	6	4.4	583	1	HEMA_INBNI	P03464 influenza b	194	6	4.4	1211	1	DP3A_HELPJ	P56157 helicobacte
122	6	4.4	584	1	HEMA_INBLE	P03463 influenza b	195	6	4.4	1225	1	Y309_MYCGE	P47351 mycoplasma
123	6	4.4	585	1	HEMA_INBBE	P17504 influenza b	196	6	4.4	1227	1	LAF4_HUMAN	P51826 homo sapien
124	6	4.4	585	1	HEMA_INBVK	P22092 influenza b	197	6	4.4	1486	1	MURK_ECOLI	P22523 escherichia
125	6	4.4	588	1	PEX5_YARLI	O99144 yarrowia li	198	6	4.4	1762	1	DPQO_HUMAN	O75417 homo sapien
126	6	4.4	627	1	YHAB_YEAST	P38750 saccharomyc	199	6	4.4	1871	1	PLX4_HUMAN	P51805 homo sapien
127	6	4.4	627	1	YHAB_YEAST	P38750 saccharomyc	200	6	4.4	1947	1	MYSC_CAEEL	P19845 caenorhabdi
128	6	4.4	628	1	HNFA_MOUSE	P22361 mus musculu	201	6	4.4	3079	1	IRR2_YEAST	P19158 saccharomyc
129	6	4.4	632	1	HNFA_MOUSE	P15257 rattus norv	202	6	4.4	3519	1	OLB6_STRAT	O07017 streptomyc
130	6	4.4	632	1	ETPD_SCHPO	P87111 schizosacch	203	6	4.4	4369	1	RYNC_RABIT	P30957 oryctolagus
131	6	4.4	638	1	KRAF_XENLA	P09360 xenopus lae	204	6	4.4	4699	1	BOYR_BOTJA	P22028 bothrops ja
132	6	4.4	651	1	HS7D_DROME	P11147 drosophila	205	6	3.7	25	1	PA2_RHONO	P43318 rhopilema n
133	6	4.4	651	1	SYM_METJA	O58659 methanococc	206	6	3.7	33	1	RLZ1_XENLA	P49628 xenopus lae
134	6	4.4	662	1	TLPB_BACSU	P39217 bacillus su	207	6	3.7	32	1	TYBN_HUMAN	O99406 homo sapien
135	6	4.4	664	1	SYM_BACSU	P37465 bacillus su	208	6	3.7	49	1	THP1_BOVIN	P01249 bos taurus
136	6	4.4	668	1	YKCA_CAEEL	P42083 caenorhabdi	209	6	3.7	49	1	THP2_BOVIN	P01250 bos taurus
137	6	4.4	688	1	YKCA_AGRT6	P09817 agrobacteri	210	6	3.7	49	1	THPS_BOVIN	P01251 bos taurus
138	6	4.4	696	1	Y195_SYNY3	P74101 synechocyst	211	6	3.7	52	1	RUBR_DESGI	P00270 desulfovibr
139	6	4.4	705	1	ADBG_RAT	O62847 rattus norv	212	6	3.7	53	1	IBB2_WHEAT	O98864 triticum ae
140	6	4.4	705	1	ADBG_RAT	O62847 rattus norv	213	6	3.7	54	1	RL33_XYLEFA	O99621 xyella fas
141	6	4.4	705	1	ADBG_RAT	O62847 rattus norv	214	6	3.7	54	1	STP1_MOUSE	P10856 mus musculu
142	6	4.4	706	1	ADBG_RAT	O62847 rattus norv	215	6	3.7	54	1	STP1_MOUSE	P03317 rattus norv
143	6	4.4	706	1	ADBG_RAT	O62847 rattus norv	216	6	3.7	54	1	Y01K_BP74	P33228 bacterioph
144	6	4.4	710	1	ADBG_MOUSE	O9y988 homo sapien	217	6	3.7	59	1	MPPE_SULAC	P23039 sulfolobus
145	6	4.4	715	1	RNR_THEMA	O9y988 mus musculu	218	6	3.7	61	1	RLZ1_PIG	P49666 sus scrofa
146	6	4.4	721	1	DNK_APHHA	O9y988 mus musculu	219	6	3.7	63	1	PERC_SUNNU	P81498 suncus murti
147	6	4.4	728	1	HGF_MOUSE	O52960 aphanothece	220	6	3.7	68	1	YVFK_ECOLI	P33324 paracoccus
148	6	4.4	728	1	HGF_MOUSE	O52960 aphanothece	221	6	3.7	71	1	HPIS_PARSP	P00264 paracoccus
149	6	4.4	744	1	GYRA_AQUAE	P17945 rattus norv	222	6	3.7	76	1	RPOZ_AQUAE	O66570 aquilex aeo
150	6	4.4	748	1	MURA_MOUSE	O67108 aquifex aeo	223	6	3.7	79	1	Y607_METJA	O56024 methanococc
151	6	4.4	750	1	MURA_MOUSE	P16332 mus musculu	224	6	3.7	80	1	FER1_METJA	O60368 methanococc
152	6	4.4	750	1	MURA_MOUSE	P22033 homo sapien	225	6	3.7	82	1	YB1M_HERAV	P25237 herpetosiph
153	6	4.4	761	1	STAT_DROME	O24151 drosophila	226	6	3.7	82	1	YELM_HERAV	P25237 herpetosiph
154	6	4.4	769	1	VPI_BPPH6	P11126 bacterioph	227	6	3.7	83	1	YELM_HERAV	P25237 herpetosiph
155	6	4.4	770	1	RNR_SHIFL	P30851 shigella fl	228	6	3.7	85	1	CAID_BOVIN	O21976 bos taurus
156	6	4.4	794	1	YE14_YEAST	P39961 saccharomyc	229	6	3.7	85	1	ACTO_ENTHI	O15602 entamoeba h
157	6	4.4	798	1	YC00_MYCPN	O63418 rattus norv	230	6	3.7	87	1	GRG_IOLFO	O01821 loligo forb
158	6	4.4	800	1	P111_YEAST	O50288 mycoplasma	231	6	3.7	88	1	GLR2_SYNY3	P73492 synechocyst
159	6	4.4	803	1	GYRB_BUCAL	P08468 saccharomyc	232	6	3.7	89	1	DBHB_NEIMA	O91416 nisseria m
160	6	4.4	813	1	RNR_ECOLI	P57126 buchnera ap	233	6	3.7	89	1	TATA_HAEIN	P57045 hemophilus
161	6	4.4	814	1	SEFC_SALEN	P21499 escherichia	234	6	3.7	91	1	Y016_METJA	O56326 methanococc
162	6	4.4	826	1	LOM2_MYXXA	P33388 salmonella	235	6	3.7	92	1	PEIC_ECOLI	P14039 methanococc
163	6	4.4	826	1	LOM2_MYXXA	P36774 myxococcus	236	6	3.7	92	1	PC01_BP22	P39159 escherichia
164	6	4.4	834	1	RSG5_HUMAN	O43374 homo sapien	237	6	3.7	92	1	VC02_SPVKA	P32230 swinepox vi
165	6	4.4	836	1	NNA3_HUMAN	P48764 homo sapien	238	6	3.7	93	1	HMCI_METPA	O9y416 methanocarc
166	6	4.4	862	1	CLPB_HELPY	P71404 helicobacte	239	6	3.7	94	1	RL31_PYRAB	O9y416 methanocarc
167	6	4.4	875	1	LOX1_HORVU	P29114 hordium vul	240	6	3.7	94	1	RL35_PIG	O9y416 methanocarc
168	6	4.4	876	1	IMB1_RAT	P53296 rattus norv	241	6	3.7	94	1	Y02W_HERAV	O29361 sus scrota
169	6	4.4	876	1	IMB1_HUMAN	O14974 homo sapien	242	6	3.7	95	1	ATP2_YEAST	P32148 herpetosiph
170	6	4.4	901	1	IMB1_MOUSE	P70168 mus musculu	243	6	3.7	95	1	DBH_THERH	P19436 thermus aqu
171	6	4.4	910	1	PHSG_YEAST	P06738 saccharomyc	244	6	3.7	96	1	CTCI_ACTIM	P19436 thermus aqu
172	6	4.4	936	1	SC15_YEAST	P22224 saccharomyc	245	6	3.7	96	1	FER_SYNIT	O33947 actinobact
173	6	4.4	937	1	CDAY_HUMAN	O9y517 homo sapien	246	6	3.7	96	1	XTLB_ARTS7	P00255 synechococc
174	6	4.4	941	1	CDAC_HUMAN	O9un72 homo sapien	247	6	3.7	96	1	TGGU_ECOLI	P26909 actinobacte
175	6	4.4	941	1	CDAC_HUMAN	O9un75 homo sapien	248	6	3.7	97	1	FER_SYNEL	P22060 escherichia
176	6	4.4	948	1	CDR2_HUMAN	O9un74 homo sapien	249	6	3.7	98	1	FER2_RAPSA	P14937 rapanus sa
177	6	4.4	948	1	CDAA_HUMAN	O9y519 homo sapien	250	6	3.7	98	1	KRFT_LARNO	P02451 larus novae
178	6	4.4	949	1	CDAB_HUMAN	O9y512 homo sapien	251	6	3.7	99	1	SRCY_BACST	P28650 bacillus st
179	6	4.4	949	1	IF2_HELPJ	O92m46 helicobacte	252	6	3.7	100	1	POL_SIVA2	P12500 simian immu



253	5	3.7	100	1	POL_STV3	P12501	stlmian	immu	326	5	3.7	128	1	RL26_CHICK	P47832	gallus	gall
254	5	3.7	100	1	RL23_BUCAI	P57589	buchnera	ap	327	5	3.7	129	1	GLNR_BACE	P19083	baecillus	ce
255	5	3.7	101	1	B3AR_PIG	O95252	sus	scrofa	328	5	3.7	129	1	RS6E_ARCFU	O29739	archaeoglob	
256	5	3.7	101	1	SMD3_YEAST	P43321	saccharomyc		329	5	3.7	129	1	VE11_VACCV	P21051	vaccinia	vi
257	5	3.7	101	1	URE2_HAEIN	P43392	haemophilus		330	5	3.7	129	1	VE11_VARY	P33822	variola	vir
258	5	3.7	101	1	YD31_MYCTU	O10642	mycobacteri		331	5	3.7	130	1	TYB8_MOUSE	P06321	mus	musculu
259	5	3.7	102	1	ARSC_NEIGO	P93354	neisseria	g	332	5	3.7	130	1	BYB12_MYCPN	P75450	mycoplasma	
260	5	3.7	102	1	CYC3_DESDE	P00134	desulfovibr		333	5	3.7	131	1	PER_ZAPVU	O21135	zapionus	t
261	5	3.7	102	1	GLRX_RICCO	P55143	ricinus	com	334	5	3.7	131	1	RYHP_SCHPO	O42984	schizosacch	
262	5	3.7	104	1	CYC_ATESP	P00003	ateles	sp..	335	5	3.7	131	1	YOHF_BACSU	P02022	rana	catesb
263	5	3.7	104	1	CYC_HUMAN	P00001	homo	saplen	336	5	3.7	132	1	HBAM_RANCA	O96716	schizosacch	
264	5	3.7	104	1	CYC_MACMU	P00002	macaca	mula	337	5	3.7	132	1	RL7B_SCHPO	O42106	narke	japon
265	5	3.7	104	1	RPOZ_STRPY	P82577	streptococc		338	5	3.7	133	1	CLX1_NARJA	P51260	porphyra	pu
266	5	3.7	105	1	ARSR_BACSU	P45949	bacillus	su	339	5	3.7	134	1	ATPE_PORFU	O18180	homo	saplen
267	5	3.7	105	1	COTW_BACSU	O08310	bacillus	su	340	5	3.7	134	1	CLX1_HUMAN	O64276	mus	musculu
268	5	3.7	105	1	Y613_ARCFU	O29642	archaeoglob		341	5	3.7	134	1	CLX1_MOUSE	O13329	homo	saplen
269	5	3.7	106	1	BOLA_VIBCH	O9KPS0	vibrio	chol	342	5	3.7	134	1	CLX2_HUMAN	O42105	narke	japon
270	5	3.7	106	1	RNFI_GIBRU	P10282	glibberella		343	5	3.7	134	1	CLX2_NARJA	P53716	candida	alb
271	5	3.7	106	1	Y094_HAEIN	P45939	haemophilus		344	5	3.7	134	1	YB5_CANLU	P06285	marchantia	
272	5	3.7	106	1	YAVC_RHISN	O53211	rhizobium	s	345	5	3.7	135	1	ATPE_MARPO	O01469	homo	saplen
273	5	3.7	106	1	YBC8_YEAST	P38202	saccharomyc		346	5	3.7	135	1	FABE_HUMAN	O08450	clostridium	
274	5	3.7	107	1	RLA1_LEIPE	O46313	leishmania		347	5	3.7	136	1	DEF_CLOBE	P01759	mus	musculu
275	5	3.7	108	1	Y203_METYA	O60268	methanococc		348	5	3.7	136	1	RL15_MOUSE	P55837	actinobacil	
276	5	3.7	108	1	FEL2_FELCA	P30440	felis	sive	349	5	3.7	136	1	RL16_ACTAC	P02414	eschericichia	
277	5	3.7	109	1	LY1F_HUMAN	P04208	homo	saplen	350	5	3.7	136	1	RL16_ECOLI	P02414	eschericichia	
278	5	3.7	109	1	RS17_METYA	P14042	methanococc		351	5	3.7	136	1	RL16_HAEIN	P44354	haemophilus	
279	5	3.7	109	1	RH10_EMENI	P29429	emeritella		352	5	3.7	136	1	Y441_MYCPN	P75149	mycoplasma	
280	5	3.7	109	1	VG1I_HSVGB	P36343	simian	hezp	353	5	3.7	136	1	Y452_CAEEL	O62250	caenorhabdi	
281	5	3.7	109	1	YF39_METYA	O58934	methanococc		354	5	3.7	137	1	HY11_MOUSE	P01755	mus	musculu
282	5	3.7	110	1	EMS2_LYTVA	P28773	lytechinus		355	5	3.7	137	1	HY46_MOUSE	P01822	mus	musculu
283	5	3.7	110	1	RLA4_YEAST	P02400	saccharomyc		356	5	3.7	137	1	Y074_MYCGE	P47320	mycoplasma	
284	5	3.7	110	1	YCKD_BACSU	P42402	bacillus	su	357	5	3.7	137	1	Y186_BORBU	O51204	botreilia	bu
285	5	3.7	111	1	HY35_MOUSE	P01804	mus	musculu	358	5	3.7	138	1	PSBU_SYN2	P17475	synechococc	
286	5	3.7	111	1	RLA1_CAEEL	P91913	caenorhabdi		359	5	3.7	139	1	HY07_MOUSE	P01751	mus	musculu
287	5	3.7	111	1	YC91_MYCTU	O10617	mycobacteri		360	5	3.7	139	1	Y024_METYA	O60334	methanococc	
288	5	3.7	112	1	ACPM_SCHPO	O10217	schizosacch		361	5	3.7	140	1	H2AV_DROME	P08985	drosophila	
289	5	3.7	112	1	PEDB_PEDAC	P36496	pediococcus		362	5	3.7	140	1	Y431_VARY	P33848	varitola	yir
290	5	3.7	112	1	RLA1_DROME	P08570	drosophila		363	5	3.7	140	1	Y649_ARCTU	O29019	archaeoglob	
291	5	3.7	113	1	FLIT_BACSU	P39740	bacillus	su	364	5	3.7	142	1	B1R5_HUMAN	O15392	homo	saplen
292	5	3.7	114	1	PSPP_PAPAN	O28767	papio	anubi	365	5	3.7	142	1	Y742_CHLNP	O92793	chlamydia	p
293	5	3.7	115	1	RL19_BACSU	O31742	bacillus	su	366	5	3.7	142	1	YECG_ECOLI	P46888	eschericichia	
294	5	3.7	115	1	RL24_SYNY3	P73309	synechocyst		367	5	3.7	143	1	DHSB_RAT	P21913	rattus	norv
295	5	3.7	117	1	PRD1_CAEEL	O17827	caenorhabdi		368	5	3.7	143	1	PRD2_DROME	O9745	drosophila	
296	5	3.7	117	1	PRDB_PYRHO	O58268	pyrococcus		369	5	3.7	143	1	RL15_METYA	P14032	methanococc	
297	5	3.7	117	1	SEPT4_HUMAN	O16550	homo	saplen	370	5	3.7	144	1	DRRI_HUMAN	O95990	homo	saplen
298	5	3.7	117	1	YCJD_ECOLI	P45736	eschericichia		371	5	3.7	144	1	NAPV_HELPY	O09590	homo	saplen
299	5	3.7	118	1	HV39_MOUSE	P01809	mus	musculu	372	5	3.7	144	1	FLAV_HELPY	P43313	hellicobacte	
300	5	3.7	118	1	RLA3_ORYSA	P56724	oryza	sativ	373	5	3.7	144	1	RPOE_MYCPN	P75090	mycoplasma	
301	5	3.7	118	1	YG35_BPLIH	O04769	lactococcus		374	5	3.7	144	1	SODM_ORYSA	P24120	oryza	sativ
302	5	3.7	119	1	WN1A_PLETO	P28131	plethodon	j	375	5	3.7	144	1	SODM_BRARL	P28761	branchiosto	
303	5	3.7	119	1	YXS9_CAEEL	O10026	caenorhabdi		376	5	3.7	144	1	Y100_AOUAE	O67669	aquilex	aeo
304	5	3.7	120	1	HV03_MOUSE	P01747	mus	musculu	377	5	3.7	145	1	RL26_HUMAN	P12749	homo	saplen
305	5	3.7	121	1	H2B1_TETTH	P08993	tetrahymena		378	5	3.7	145	1	RL26_RAT	O33603	streptococc	
306	5	3.7	121	1	H2B2_TETTH	P08994	tetrahymena		379	5	3.7	145	1	SODM_STRAI	O54210	streptococc	
307	5	3.7	121	1	HY01_MOUSE	P07745	mus	musculu	380	5	3.7	145	1	SODM_STROB	O54233	streptococc	
308	5	3.7	122	1	RL35_HUMAN	P42766	homo	saplen	381	5	3.7	145	1	SODM_STRIN	O54266	streptococc	
309	5	3.7	122	1	RL35_RAT	P17078	rattus	norv	382	5	3.7	145	1	SODM_STRO	O54286	streptococc	
310	5	3.7	122	1	SSB_BPB03	O37885	bacteriophya		383	5	3.7	145	1	SODM_STRSL	O33733	streptococc	
311	5	3.7	122	1	YEBU_ECOLI	P76364	eschericichia		384	5	3.7	146	1	FLAV_DESGI	O01095	desulfovibr	
312	5	3.7	123	1	YB8F_YEAST	P38357	saccharomyc		385	5	3.7	146	1	RL13_MYCGE	P47657	mycoplasma	
313	5	3.7	123	1	YGK6_YEAST	P53077	saccharomyc		386	5	3.7	146	1	RL26_BRARA	O39411	brassica	ra
314	5	3.7	123	1	Y0S4_YEAST	P46984	saccharomyc		387	5	3.7	147	1	YBPF_BACSU	P50732	bacillus	su
315	5	3.7	123	1	YROP_HALHA	O24785	halobacteri		388	5	3.7	148	1	CADF_DROME	P45594	drosophila	
316	5	3.7	124	1	CHEY_PSEAE	O51455	pseudomonas		389	5	3.7	148	1	D112_MOUSE	O091m83	mus	musculu
317	5	3.7	124	1	MUTY_BUCAI	P57298	buchnera	ap	390	5	3.7	148	1	DSBH_BPSPC	O64037	bacteriophya	
318	5	3.7	124	1	PER_HTRPI	O25109	hirtodrosop		391	5	3.7	149	1	LSM4_FAGSY	O94929	fagus	sylya
319	5	3.7	124	1	VA31_VACCC	P21096	vaccinia	vi	392	5	3.7	149	1	HMGL_VICFA	P40620	vicia	faba
320	5	3.7	124	1	VA31_VACCV	P24760	vaccinia	vi	393	5	3.7	149	1	MP17_FFRATU	P18149	francisella	
321	5	3.7	124	1	YEF7_SCHPO	O14068	schizosacch		394	5	3.7	149	1	RL15_AOUAE	O67561	aquilex	aeo
322	5	3.7	124	1	YHCC_BACSU	P54587	bacillus	su	395	5	3.7	149	1	RL15_METYA	O58429	methanococc	
323	5	3.7	125	1	CG80_HUMAN	O9y3b4	homo	saplen	396	5	3.7	150	1	DUT_LACIA	O9c730	lactococcus	
324	5	3.7	126	1	RS6E_METTH	O26360	methanobact		397	5	3.7	150	1	KPSH_HUMAN	P11801	homo	saplen
325	5	3.7	127	1	SYUR_HUMAN	O9n250	homo	saplen	398	5	3.7	151	1	INO4_YEAST	P13902	saccharomyc	

399	5	3.7	151	1	RL13_SYNY3	P73294 synechocyst	472	5	3.7	175	1	YF49_MYCTU	Q10777 mycobacteri
400	5	3.7	152	1	MSCL_CLOPE	P53380 clostridium	473	5	3.7	176	1	TRAF_AGR16	Q44364 agrobacteri
401	5	3.7	152	1	Y554_METUA	057974 methanococc	474	5	3.7	177	1	BTC_MOUSE	Q05928 mus musculi
402	5	3.7	152	1	YB83_METUA	058878 methanococc	475	5	3.7	177	1	RL6_METTH	Q26127 methanobact
403	5	3.7	153	1	Y264_METUA	057712 methanococc	476	5	3.7	177	1	YAI6_SCHPO	Q09688 schizosacch
404	5	3.7	153	1	YEH3_PSEAE	Q9hb0 pseudomonas	477	5	3.7	177	1	YNIW_AZCOH	P23177 azotobacter
405	5	3.7	154	1	BFRA_NEIMA	P56998 neisseria m	478	5	3.7	178	1	BTC_HUMAN	P35070 homo sapien
406	5	3.7	154	1	BFRA_NEIMA	P72080 neisseria m	479	5	3.7	178	1	RBS_TRIRP	P17673 trifolium r
407	5	3.7	154	1	RR7_SPIOL	P81229 splinacia ol	480	5	3.7	178	1	SIGY_BACSU	P94370 bacillus su
408	5	3.7	154	1	YK01_AERPE	Q9vnd8 aeropyrum p	481	5	3.7	178	1	VG20_BPLH	Q04763 lactococcus
409	5	3.7	154	1	YMI3_MARPO	09vnd8 aeropyrum p	482	5	3.7	179	1	MUSG_STRPY	P82547 streptococc
410	5	3.7	154	1	YME4_BACSU	P38457 marchantia	483	5	3.7	179	1	Y502_STRO	Q9x928 streptomyc
411	5	3.7	155	1	RL21_PYUST	P39632 bacillus su	484	5	3.7	179	1	YAI1_ECOLI	P46122 escherichia
412	5	3.7	155	1	RR7_ARATH	P49667 pyura stolo	485	5	3.7	180	1	HMG1_CRIGR	P07156 cricetus
413	5	3.7	155	1	RR7_CUSEU	P56800 arabidopsis	486	5	3.7	180	1	RBS2_PEA	P00869 pisum sativ
414	5	3.7	155	1	RR7_CUSEU	P46292 cuscuta eur	487	5	3.7	180	1	RBS3_PEA	P07689 pisum sativ
415	5	3.7	155	1	RR7_CUSEU	P34832 cuscuta ref	488	5	3.7	180	1	RBS3_PEA	Q65194 medicago sa
416	5	3.7	155	1	RR7_EPRVI	P30057 epifagus vi	489	5	3.7	180	1	RL15_LEPIN	Q9yd17 leptospira
417	5	3.7	155	1	RR7_EUGGR	P02360 euglena gra	490	5	3.7	180	1	VG20_BPMVA	Q04764 lactococcus
418	5	3.7	155	1	RR7_MARPO	P06360 marchantia	491	5	3.7	180	1	YI6A_MYCMT	Q05239 mycoplasma
419	5	3.7	155	1	RR7_PINTH	P41652 pinus thunb	492	5	3.7	180	1	YMI2_MARPO	P38464 marchantia
420	5	3.7	155	1	RR7_SOYBN	P07135 glycine max	493	5	3.7	181	1	ATP3_HUMAN	P18847 homo sapien
421	5	3.7	155	1	RR7_TOBAC	P06361 nicotiana t	494	5	3.7	181	1	ATP3_MOUSE	Q60765 mus musculi
422	5	3.7	155	1	RS15_HAIMA	P05762 haloarcula	495	5	3.7	181	1	ATP3_MOUSE	P29596 rattus norv
423	5	3.7	156	1	SSRP_BACSU	Q32230 bacillus su	496	5	3.7	181	1	RBS1_NITCSY	P22433 nicotiana s
424	5	3.7	157	1	YI17_YEAST	P40502 saccharomyc	497	5	3.7	181	1	RL5_METVA	P14029 methanococc
425	5	3.7	158	1	FMA4_BACNO	P02975 bacteroides	498	5	3.7	181	1	RM06_ACACA	P46765 acanthamoeb
426	5	3.7	158	1	ILVH_LACLA	Q02140 lactococcus	499	5	3.7	182	1	HSLV_RICPR	Q9xdk9 rickettsia
427	5	3.7	159	1	RL21_HUMAN	P46778 homo sapien	500	5	3.7	182	1	YI1P3_CAEL	P34383 caenorhabdi
428	5	3.7	159	1	RL21_MOUSE	Q09167 mus musculi	501	5	3.7	182	1	YOPO_YEREN	P27474 versinia en
429	5	3.7	159	1	RL21_RAT	P20280 rattus norv	502	5	3.7	183	1	APR_ECOLI	P07672 escherichia
430	5	3.7	159	1	Y268_BORBU	Q44756 borrelia bu	503	5	3.7	184	1	ETS3_DROME	P29774 drosophila
431	5	3.7	160	1	FMA1_BACNO	P17823 bacteroides	504	5	3.7	185	1	CBX1_HUMAN	P23197 homo sapien
432	5	3.7	160	1	FMA2_BACNO	P17824 bacteroides	505	5	3.7	185	1	PF03_CAEL	Q18054 caenorhabdi
433	5	3.7	160	1	FMA3_BACNO	P27691 bacteroides	506	5	3.7	185	1	YALI_TRYBB	P17960 trypanosoma
434	5	3.7	160	1	FMA3_BACNO	P17822 bacteroides	507	5	3.7	186	1	CHS2_USTMA	P30599 usutillago ma
435	5	3.7	160	1	YCX3_ASTIO	P34777 astasia lon	508	5	3.7	186	1	MOS2_SCHPO	Q11118 schizosacch
436	5	3.7	160	1	YDCD_ECOLI	P31991 escherichia	509	5	3.7	187	1	ALL1_HORSE	Q95188 equus caball
437	5	3.7	160	1	YEMO_YEAST	P40019 saccharomyc	510	5	3.7	187	1	DEF_SYNY3	P73441 synechocyst
438	5	3.7	161	1	FMA3_BACNO	P27689 bacteroides	511	5	3.7	187	1	PTPA_PSEAE	Q58641 pseudomonas
439	5	3.7	161	1	FMA3_BACNO	P27906 bacteroides	512	5	3.7	188	1	DCYD_BPT2	P00814 bacterioph
440	5	3.7	162	1	VENY_EAV	P28991 equine arte	513	5	3.7	188	1	VP21_TBSV8	P50629 tomato bush
441	5	3.7	162	1	Y480_TREPA	Q83493 treponema p	514	5	3.7	188	1	YEBB_SCHPO	Q14076 schizosacch
442	5	3.7	164	1	PHAI_SYNPP	Q02179 synechococc	515	5	3.7	189	1	CHS2_XYIIBA	P30604 xylohypha b
443	5	3.7	164	1	PHAI_SYNPP	Q02179 synechococc	516	5	3.7	189	1	VP21_ANCV	P15961 artichoke m
444	5	3.7	164	1	Y859_METUA	P45072 haemophilus	517	5	3.7	189	1	VP21_TBSVA	P50630 tomato bush
445	5	3.7	164	1	Y859_METUA	Q58269 methanococc	518	5	3.7	189	1	VP21_TBSVB	P50631 tomato bush
446	5	3.7	165	1	YSEA_SPACA	P47995 staphylococ	519	5	3.7	190	1	ARE_GIALA	P54063 methanococc
447	5	3.7	165	1	UTXA_CLODI	P16153 clostridium	520	5	3.7	191	1	RS7_METUA	Q67376 aquifex aeo
448	5	3.7	168	1	BI13_CAUCR	P34008 caulobacter	521	5	3.7	192	1	ERP_AOUAE	P43408 methanococc
449	5	3.7	168	1	FMS3_ECOLI	Q03683 nicotiana t	522	5	3.7	192	1	KADA_METTG	Q58708 pyrococcus
450	5	3.7	169	1	Y358_BUCAI	P57439 buchnera ap	523	5	3.7	192	1	MOBA_PPRHO	Q58708 pyrococcus
451	5	3.7	169	1	Y631_METUA	Q58048 methanococc	524	5	3.7	192	1	PAAD_CHLMO	Q9ph2 chlamydia m
452	5	3.7	170	1	FMF7_ECOLI	P25394 escherichia	525	5	3.7	193	1	DCYD_BPT4	P31606 bacterioph
453	5	3.7	170	1	VEAR_HCMVA	P06694 human cytom	526	5	3.7	193	1	GVE2_HALN1	Q48311 halobacteri
454	5	3.7	171	1	CD3D_HUMAN	P04334 homo sapien	527	5	3.7	194	1	RS7_METVA	P14037 methanococc
455	5	3.7	171	1	RT25_MOUSE	Q9d125 mus musculi	528	5	3.7	194	1	YAI1_ARCFU	Q29261 archaeoglob
456	5	3.7	171	1	Y328_AOUAE	Q66665 aquifex aeo	529	5	3.7	194	1	YAI1_ARCFU	P37050 escherichia
457	5	3.7	171	1	Y708_TREPA	Q83706 treponema p	530	5	3.7	195	1	ALIA_YEAST	P32459 saccharomyc
458	5	3.7	171	1	YE11_HAEIN	Q57374 haemophilus	531	5	3.7	195	1	GYRA_FIBSU	P35810 fibroblacter
459	5	3.7	172	1	PHCB_CYACA	Q19909 cyanidium c	532	5	3.7	196	1	CAGS_HELPY	P97227 helicobacte
460	5	3.7	172	1	TCTP_MOUSE	P14701 mus musculi	533	5	3.7	196	1	HP20_TAMAS	Q06575 tamias asia
461	5	3.7	172	1	YB03_MYCPN	P75366 mycoplasma	534	5	3.7	196	1	RK3_ODOST	P49569 odontella s
462	5	3.7	173	1	ERG_LYTYA	Q01414 lytechinus	535	5	3.7	196	1	YPB3_LACLA	P24097 lactococcus
463	5	3.7	173	1	HA34_BRELC	Q99074 brentia lact	536	5	3.7	197	1	ENGB_ARCFU	Q28943 archaeoglob
464	5	3.7	173	1	LYR2_MOUSE	Q08731 mus musculi	537	5	3.7	197	1	RUVA_BORBU	P70827 borrelia bu
465	5	3.7	173	1	YRVS_CAUEL	Q09355 caenorhabdi	538	5	3.7	197	1	Y060_BPT4	P33924 bacterioph
466	5	3.7	174	1	FANG_ECOLI	P20861 escherichia	539	5	3.7	197	1	YF79_ARCFU	Q28693 archaeoglob
467	5	3.7	174	1	GRPE_METTH	Q27350 methanobact	540	5	3.7	198	1	RAC3_ARATH	Q38912 arabidopsis
468	5	3.7	175	1	COAG_CARRO	P03997 carcinoscor	541	5	3.7	199	1	RS8_YEAST	P05754 saccharomyc
469	5	3.7	175	1	ELBS_ADEMI	P12535 mouse adeno	542	5	3.7	199	1	SODF_BABBO	Q15905 babesia bov
470	5	3.7	175	1	IMMB_ECOLI	P22426 escherichia	543	5	3.7	199	1	YGAT_YEAST	P50083 saccharomyc
471	5	3.7	175	1	RL20_SCHPO	P05732 schizosacch	544	5	3.7	199	1	YGCH_ECOLI	Q46697 escherichia

545	5	3.7	200	1	PCR_PYRAB	Q9uy9	pyrococcus	618	5	3.7	224	1	GER2_WHEAT	P15290	triticum ae
546	5	3.7	200	1	RECR_CHLMU	Q9px4	chlamydia m	619	5	3.7	224	1	KST5_ECOLI	P2456	escherichia
547	5	3.7	200	1	RECR_CHLMU	O84243	chlamydia t	620	5	3.7	224	1	MALA_STRPN	P48186	streptococ
548	5	3.7	200	1	RECR_HAEN	P44712	haemophilus	621	5	3.7	224	1	RS3_CHLMU	O9pj0	chlamydia t
549	5	3.7	200	1	RECR_PASMU	P57826	pasteurella	622	5	3.7	224	1	RS3_CHLMU	O84527	chlamydia m
550	5	3.7	200	1	SODM_STRPY	P77957	streptococ	623	5	3.7	224	1	Y119_MENTU	O57583	methanococ
551	5	3.7	200	1	SYST_LYCES	P27058	lycopersico	624	5	3.7	224	1	YF81_MCTU	O50640	methanococ
552	5	3.7	201	1	ABP1_MAIZE	P13659	zea mays (m	625	5	3.7	225	1	GP30_BPSP1	O38423	bacterioph
553	5	3.7	202	1	CR2_HORVU	P23252	hordeum vul	626	5	3.7	225	1	KAD2_YEAST	P29346	saccharomyc
554	5	3.7	202	1	SODM_STRMU	P09738	streptococ	627	5	3.7	225	1	RS3_LEPIN	O9xd30	leptospiro
555	5	3.7	202	1	VANX_ENTFA	O47749	enterococcu	628	5	3.7	225	1	Y851_PYRAB	O58220	methanococ
556	5	3.7	203	1	CLPP_MARPO	P12208	marichantia	629	5	3.7	225	1	Y851_PYRAB	O9uy40	pyrococcus
557	5	3.7	203	1	THGA_ECOLI	P07464	escherichia	630	5	3.7	227	1	TPIS_PYRMO	P95153	pyrococcus
558	5	3.7	203	1	YBM4_YEAST	P38068	saccharomyc	631	5	3.7	227	1	UBU_DROME	P35122	dirosophila
559	5	3.7	203	1	YKVA_YEAST	P36036	saccharomyc	632	5	3.7	227	1	YKVA_METTF	P29577	methanobact
560	5	3.7	204	1	UBCX_PICPA	P49408	pichia past	633	5	3.7	228	1	GLXC_RHIME	O87391	rhizobium m
561	5	3.7	205	1	GARL_YEAST	P28007	saccharomyc	634	5	3.7	228	1	LOLD_BUCAI	P67363	buchnera ap
562	5	3.7	206	1	KAD_AQUAE	O66420	aquifex aeo	635	5	3.7	228	1	TPIS_PYRAB	O9uxx2	pyrococcus
563	5	3.7	206	1	Y817_METUA	O58227	methanococ	636	5	3.7	228	1	Y855_MENTU	O58265	methanococ
564	5	3.7	207	1	FLA2_PYRHO	O58283	pyrococcus	637	5	3.7	228	1	Y503_ACIAM	P29087	acidianus a
565	5	3.7	207	1	PSAD_CUCSA	P32869	cucumis sat	638	5	3.7	228	1	YTUB_ERWHE	O47826	erwinia her
566	5	3.7	207	1	SODM_YEREN	P53655	yersinia en	639	5	3.7	229	1	MTN_HAEN	P51113	haemophilus
567	5	3.7	208	1	ACPD_BACSU	O35022	bacillus su	640	5	3.7	229	1	PEPE_SALTY	P36936	salmonella
568	5	3.7	208	1	HRPW_PSESY	O60236	pseudomonas	641	5	3.7	230	1	Y643_PYRHO	O58377	pyrococcus
569	5	3.7	208	1	MLEY_HUMAN	P14649	homo sapien	642	5	3.7	231	1	TPIS_PYRHO	O59536	pyrococcus
570	5	3.7	210	1	HIS7_MYCLE	Q9x709	mycobacteri	643	5	3.7	231	1	UL71_HSV62	P23546	human herpe
571	5	3.7	210	1	SSH5_YEAST	Q03446	saccharomyc	644	5	3.7	232	1	CASB_PIG	P39037	sus scrofa
572	5	3.7	210	1	XYNA_BACST	P45705	bacillus st	645	5	3.7	232	1	KLK_PIG	P00754	sus scrofa
573	5	3.7	210	1	Y041_BP4	P70704	bacterioph	646	5	3.7	232	1	THIQ_ECOLI	P31548	escherichia
574	5	3.7	210	1	YEO9_YEAST	PA0052	saccharomyc	647	5	3.7	232	1	VHED_BP73	P20313	bacterioph
575	5	3.7	212	1	KAD_STRPY	P82549	streptococ	648	5	3.7	232	1	VHED_BP73	P30366	bacterioph
576	5	3.7	212	1	UL03_HSV6B	P28942	equine herp	649	5	3.7	232	1	YC68_MYCTU	Q11051	mycobacteri
577	5	3.7	213	1	CAT2_ECOLI	P22615	escherichia	650	5	3.7	233	1	CLC1_YEAST	P17891	saccharomyc
578	5	3.7	213	1	UL71_HSV6U	P52473	human herpe	651	5	3.7	233	1	YOD2_CABEL	P34594	caenorhabdi
579	5	3.7	213	1	XYNA_BACCI	P09850	human herpe	652	5	3.7	236	1	YCBF_ECOLI	PA0876	escherichia
580	5	3.7	213	1	XYNA_BACCI	P18429	bacillus ci	653	5	3.7	236	1	YIEK_ECOLI	P31470	escherichia
581	5	3.7	214	1	CYB_AGRKO	P92845	agkistrodon	654	5	3.7	237	1	RECA_NEIEG	P03396	neisseria e
582	5	3.7	214	1	CYB_AGRKA	P92852	agkistrodon	655	5	3.7	237	1	RECA_NEIEP	O86403	neisseria p
583	5	3.7	214	1	CYB_BOTAT	P92846	bothrops at	656	5	3.7	237	1	RECA_NEISI	O86411	neisseria s
584	5	3.7	214	1	CYB_BOTSC	P92847	bothriopsis	657	5	3.7	237	1	RECA_NEISU	O86416	neisseria s
585	5	3.7	214	1	CYB_BOTSC	P92849	bothriopsis	658	5	3.7	237	1	RECA_NEISU	P26864	marichantia
586	5	3.7	214	1	CYB_CROAT	P92850	crotales at	659	5	3.7	238	1	GUB_PAEPD	PA5797	penicillium
587	5	3.7	214	1	CYB_CROAT	P92851	timeresuru	660	5	3.7	238	1	PELX_ERMCA	P16530	erwinia car
588	5	3.7	214	1	HMGI_BOVIN	P10103	bos taurus	661	5	3.7	238	1	RTBB_ARCFU	O28173	archaeoglob
589	5	3.7	214	1	HMGI_HUMAN	P09429	homo sapien	662	5	3.7	238	1	RPE_YEAST	P46965	saccharomyc
590	5	3.7	214	1	HMGI_MOUSE	P07155	mus musculu	663	5	3.7	238	1	YBM9_SCHPO	Q10333	schizosacch
591	5	3.7	214	1	HMGI_MOUSE	P12682	sus scrofa	664	5	3.7	239	1	DHSB_PORPU	P80477	porphyra pu
592	5	3.7	214	1	MPP9_HUMAN	O99550	homo sapien	665	5	3.7	239	1	DHSB_RECAM	P00480	recclinomona
593	5	3.7	215	1	VFP_NPVAC	P28707	autographa	666	5	3.7	239	1	MPUL_ARATH	O91113	arabidopsis
594	5	3.7	215	1	YKL7_YEAST	P28707	saccharomyc	667	5	3.7	239	1	RL32_HALNI	O9hp47	halobacteri
595	5	3.7	216	1	AKA2_ARATH	P28185	arabidopsis	668	5	3.7	239	1	RP35_BACTK	P26763	bacillus th
596	5	3.7	216	1	YSP4_CABEL	Q10912	caenorhabdi	669	5	3.7	239	1	RPSE_BACSU	P06222	bacillus su
597	5	3.7	217	1	GTN1_SOLTU	P32111	solanum tub	670	5	3.7	239	1	TONB_ECOLI	P02929	escherichia
598	5	3.7	218	1	DNEI_CHLHU	O32001	chlamydomon	671	5	3.7	240	1	OM31_BRUME	Q45312	bruceella me
599	5	3.7	218	1	Y052_BORBU	Q51081	borrelia bu	672	5	3.7	240	1	YFIC_HAEN	PA47002	haemophilus
600	5	3.7	218	1	YF81_MYCLE	Q49649	mycobacteri	673	5	3.7	241	1	RECO_SALTY	O56056	salmonella
601	5	3.7	219	1	Y06K_BP4	P13318	bacterioph	674	5	3.7	242	1	CRTW_ALCSP	Q44261	alcaligenes
602	5	3.7	219	1	Y01Y_BACSU	P54536	bacillus su	675	5	3.7	242	1	RECO_ECOLI	P15027	escherichia
603	5	3.7	220	1	HEP1_BACST	P55774	bacillus st	676	5	3.7	242	1	TONB_SALTY	P25945	salmonella
604	5	3.7	220	1	VANR_ENTFA	O47744	enterococcu	677	5	3.7	242	1	TRP2_CVACA	P44793	cyanidium c
605	5	3.7	221	1	GTAL_BOVIN	O28035	bos taurus	678	5	3.7	243	1	TM17_ARATH	O59375	halobacteri
606	5	3.7	221	1	GTAL_PIG	P51781	sus scrofa	679	5	3.7	243	1	Y542_CHLPN	Q52810	chlamydia s
607	5	3.7	221	1	PLL2_MESAU	P14059	mesocricetu	680	5	3.7	246	1	MCT4_RAT	P97592	rattus norv
608	5	3.7	221	1	PLL2_RAT	P09331	rattus norv	681	5	3.7	246	1	Y586_BUCAI	P57646	buchnera ap
609	5	3.7	221	1	SEGA_BP4	P32266	bacterioph	682	5	3.7	246	1	YD53_SCHPO	O14119	schizosacch
610	5	3.7	221	1	Y532_AQUAE	O66814	aquifex aeo	683	5	3.7	247	1	PCNA_HALNI	O9h455	halobacteri
611	5	3.7	221	1	YFHB_YEAST	P43552	saccharomyc	684	5	3.7	248	1	PHON_PROST	P26975	providencia
612	5	3.7	222	1	PARA_AGRTU	P07175	agrobacteri	685	5	3.7	248	1	TRPA_PYRAB	O9vi98	pyrococcus
613	5	3.7	223	1	GTX1_TOBAC	Q03662	nicotiana t	686	5	3.7	249	1	BA72_EURSP	P19337	eubacterium
614	5	3.7	223	1	GTX2_TOBAC	Q03662	nicotiana t	687	5	3.7	249	1	PCNA_PYRAB	Q9uy48	pyrococcus
615	5	3.7	223	1	GTX3_TOBAC	Q03664	nicotiana t	688	5	3.7	249	1	PCNA_PYRHO	O73947	pyrococcus
616	5	3.7	223	1	RS3_CHLPN	O92733	chlamydia p	689	5	3.7	249	1	PCNA_PYRHO	O58338	pyrococcus
617	5	3.7	223	1	SAMP_CANPO	PA9255	cavia porce	690	5	3.7	249	1	PCNA_THERPM	Q9uw29	thermococcu

691	5	3.7	249	1	PHOC_MORMO	P28581 moranella	764	5	3.7	271	1	BLAC_PROVU	P80298 proteus vul
692	5	3.7	250	1	DHSB_CHOIR	P48932 chondirus cr	765	5	3.7	272	1	ERG_MOUSE	P81270 mus musculu
693	5	3.7	250	1	FCEB_MOUSE	P20489 mus musculu	766	5	3.7	272	1	PHEA_METJA	O58054 methanococ
694	5	3.7	250	1	YK67_CAEEL	P34345 caenorhabdi	767	5	3.7	272	1	YKEA_BACSU	O00777 bacillus su
695	5	3.7	251	1	CB24_ARATH	P27521 arabidopsis	768	5	3.7	272	1	YKNS_CAEEL	P32743 caenorhabdi
696	5	3.7	251	1	CRB1_HUMAN	P53674 homo sapien	769	5	3.7	273	1	FOLD_MYCGE	P47259 mycoplasma
697	5	3.7	251	1	IOUR_BACSU	P46337 bacillus su	770	5	3.7	273	1	SC65_YEAST	P29478 saccharomyc
698	5	3.7	252	1	I433_DICDI	P54632 dictyostell	771	5	3.7	274	1	DAPD_BUCAT	P57323 buchnera ap
699	5	3.7	252	1	PSM1_HALVO	O94V66 halobacteri	772	5	3.7	274	1	MEPA_ECOLI	P14007 escherichia
700	5	3.7	252	1	PSM1_HALVO	O94V66 halobacteri	773	5	3.7	274	1	RECA_NEIFL	O50597 neisseria f
701	5	3.7	252	1	Y830_METJA	O58418 methanococ	774	5	3.7	274	1	RECA_NEIFL	O50597 neisseria f
702	5	3.7	252	1	Y830_METJA	O58418 methanococ	775	5	3.7	274	1	RECA_NEIFL	O50597 neisseria f
703	5	3.7	253	1	YCAL_PLAFA	P25407 plasmodium	776	5	3.7	274	1	YFDO_ECOLI	O50597 neisseria f
704	5	3.7	253	1	YKXW_BACSU	P40949 bacillus su	777	5	3.7	275	1	YFDO_ECOLI	O50597 neisseria f
705	5	3.7	254	1	ADH_DHOWI	O05111 caenorhabdi	778	5	3.7	275	1	KDGA_RICPR	O39684 rickettsia
706	5	3.7	254	1	YVON_BACSU	O06977 bacillus su	779	5	3.7	275	1	PORE_SOLTU	P53399 saccharomyc
707	5	3.7	255	1	I43P_ARATH	P42644 arabidopsis	780	5	3.7	275	1	VGIM_PUMWB	P41264 pumala vir
708	5	3.7	255	1	ATP6_TRIRU	O36835 trichophyro	781	5	3.7	276	1	Y578_METJA	O57998 methanococ
709	5	3.7	255	1	HCD2_DROME	O18404 drosophila	782	5	3.7	277	1	KNOB_PLAED	P03729 plasmodium
710	5	3.7	255	1	IHRB_MOUSE	O04399 mus musculu	783	5	3.7	277	1	RS2_CHLPP	O927K9 porphyra pu
711	5	3.7	255	1	LACR_LACLA	P18816 lactococcus	784	5	3.7	277	1	THIG_PORPU	P51361 ratiuss norv
712	5	3.7	255	1	MYPS_BOVIN	P17667 bos taurus	785	5	3.7	278	1	FASL_RAT	O06667 human cytom
713	5	3.7	255	1	RECO_BACSU	P42095 bacillus su	786	5	3.7	278	1	VGIL_HCMV1	O06667 human cytom
714	5	3.7	255	1	Y439_PYRHO	O58186 pyrococcus	787	5	3.7	278	1	VGIL_HCMV1	O06667 human cytom
715	5	3.7	256	1	NORE_CHLPP	O928B3 chlamydia p	788	5	3.7	278	1	VGIL_HCMV2	O06667 human cytom
716	5	3.7	256	1	TPRS_VIBMA	P50921 vibrio mali	789	5	3.7	278	1	VGIL_HCMV3	O06667 human cytom
717	5	3.7	256	1	YGBI_HAETN	P44978 haemophilus	790	5	3.7	278	1	VGIL_HCMV4	O06667 human cytom
718	5	3.7	257	1	Y418_METJA	O57861 methanococ	791	5	3.7	278	1	VGIL_HCMV5	O06667 human cytom
719	5	3.7	258	1	ADH_BACOL	O94ar7 bacterocera	792	5	3.7	278	1	VGIL_HCMV6	O06667 human cytom
720	5	3.7	258	1	HIS6_BUCAT	P80869 bacillus su	793	5	3.7	278	1	VGIL_HCMV7	O06667 human cytom
721	5	3.7	258	1	HIS6_BUCAT	P57206 buchnera ap	794	5	3.7	278	1	VGIL_HCMV8	O06667 human cytom
722	5	3.7	258	1	NADC_ARCPR	O28439 archaeoglob	795	5	3.7	278	1	VGIL_HCMV9	O06667 human cytom
723	5	3.7	258	1	RM24_YEAST	P36525 saccharomyc	796	5	3.7	279	1	FASL_MOUSE	O06667 human cytom
724	5	3.7	258	1	YULB_BACSU	O05261 bacillus su	797	5	3.7	279	1	GR78_PLAFA	P12794 plasmodium
725	5	3.7	259	1	NAHB_PSEPU	O52459 pseudomonas	798	5	3.7	280	1	DHSB_HUMAN	P28912 homo sapien
726	5	3.7	260	1	DH10_ARATH	P42759 arabidopsis	799	5	3.7	281	1	AAC9_MICCH	P28910 micromonosp
727	5	3.7	260	1	GRAA_MOUSE	P11032 mus musculu	800	5	3.7	281	1	FASL_HUMAN	P48023 homo sapien
728	5	3.7	260	1	PDXJ_XYLFA	O9ph84 xyliella fas	801	5	3.7	281	1	NAE1_THEMA	O94967 chernobacte
729	5	3.7	261	1	YOGK_BACSU	P46342 bacillus su	802	5	3.7	281	1	PA4H_CHYRO	P30967 chromobacte
730	5	3.7	261	1	143D_SOYBN	O96453 glycine max	803	5	3.7	281	1	THTR_SACPR	P13106 saccharopoi
731	5	3.7	261	1	H1S6_SYNY3	P74106 synechocyst	804	5	3.7	281	1	XYLF_PSEPU	P23106 pseudomonas
732	5	3.7	261	1	KLK8_RAT	P36374 rattus norv	805	5	3.7	282	1	NADA_APLCA	P28241 aplysia cal
733	5	3.7	261	1	OMPY_CHLMU	O9PK23 chlamydia m	806	5	3.7	282	1	ROC_XENLA	P19600 xenopus lae
734	5	3.7	261	1	OMPY_CHLMU	O9PK23 chlamydia m	807	5	3.7	282	1	Y32K_BNYG	P19231 beet necrot
735	5	3.7	261	1	TNFS_HUMAN	P28965 homo sapien	808	5	3.7	283	1	ARBG_ERWCH	P26211 methanopyru
736	5	3.7	261	1	Y046_METJA	O60354 methanococ	809	5	3.7	283	1	MTD_METKA	P33521 matz strip
737	5	3.7	262	1	APAI_BRARE	O42363 brachydantio	810	5	3.7	283	1	VNS4_MSTV	P57649 buchera ap
738	5	3.7	262	1	GRAA_HUMAN	P12544 homo sapien	811	5	3.7	284	1	DAPF_BUCAT	O50793 mycobacteri
739	5	3.7	262	1	ISPA_AOUAE	O66952 aquifex aeo	812	5	3.7	284	1	ERR_MYCTU	P28844 pseudomonas
740	5	3.7	262	1	LPXA_HAETN	P43887 haemophilus	813	5	3.7	284	1	LEP_PSEPU	O28844 pseudomonas
741	5	3.7	262	1	LPXA_PASWU	O9cjk8 pasteurella	814	5	3.7	284	1	RP32_BUCAT	O05385 buchnera ap
742	5	3.7	262	1	MODE_ECOLI	P46930 escherichia	815	5	3.7	284	1	TPM1_SCHNA	O42637 schistosoma
743	5	3.7	262	1	PLC1_CAEEL	O93841 caenorhabdi	816	5	3.7	284	1	DVRB_ZYMON	O66998 zymomonas m
744	5	3.7	262	1	RNC_MYCGE	P47607 mycoplasma	817	5	3.7	285	1	ALKE_BABBO	O40690 treponema m
745	5	3.7	262	1	S3AD_ECOLI	P08881 klebsiella	818	5	3.7	285	1	FLA3_TREMA	O40690 treponema m
746	5	3.7	263	1	S3AD_ECOLI	P04826 escherichia	819	5	3.7	285	1	FLA3_TREMA	O40690 treponema m
747	5	3.7	264	1	CSH_ARTSP	P32400 arthróbacte	820	5	3.7	285	1	FLB2_TREHY	P21992 treponema p
748	5	3.7	264	1	GIDB_SYNY3	O55787 synechocyst	821	5	3.7	286	1	AGAY_ECOLI	P21992 treponema p
749	5	3.7	264	1	TRPA_SYNY3	P77960 synechocyst	822	5	3.7	286	1	FDHD_WOLSU	P42908 escherichia
750	5	3.7	265	1	PANB_AOUAE	O67783 aquifex aeo	823	5	3.7	286	1	FLA1_TREPA	P21818 wolliella s
751	5	3.7	265	1	YEEL_ECOLI	P76346 escherichia	824	5	3.7	286	1	FLA2_TREMA	P21990 treponema p
752	5	3.7	266	1	KLR6_MOUSE	O60653 mus musculu	825	5	3.7	286	1	FLA2_TREMA	P21990 treponema p
753	5	3.7	266	1	YF11_PYRHO	O59180 pyrococcus	826	5	3.7	286	1	FLA2_TREMA	P21990 treponema p
754	5	3.7	266	1	YRHG_BACSU	O05399 bacillus su	827	5	3.7	286	1	YAY8_SCHPO	P44749 h. dimethyla
755	5	3.7	267	1	KDGA_AOUAE	O6496 aquifex aeo	828	5	3.7	287	1	HUS1_SCHPO	P44749 h. dimethyla
756	5	3.7	268	1	ETIB_XENLA	P18756 xenopus lae	829	5	3.7	287	1	ESGA_HAETN	P44749 h. dimethyla
757	5	3.7	268	1	HISJ_NEIGO	O06758 neisseria g	830	5	3.7	288	1	ESGA_HAETN	P44749 h. dimethyla
758	5	3.7	268	1	ILIA_CAPII	P79161 capra hircu	831	5	3.7	288	1	Y134_MYCTU	O50599 mycobacteri
759	5	3.7	268	1	KNH1_YEAST	P50112 saccharomyc	832	5	3.7	288	1	EPMO_MOUSE	O00262 mus musculu
760	5	3.7	268	1	THIM_ARCPR	O28204 archaeoglob	833	5	3.7	289	1	HEM3_MERTH	O26960 methanobact
761	5	3.7	268	1	YJ80_AERPE	O94a19 aeropyrum p	834	5	3.7	289	1	SC13_PICPA	P33024 pichia past
762	5	3.7	269	1	YHIC_LACLA	O01150 lactococcus	835	5	3.7	289	1	SGCD_MESAU	P97281 mesocricetu
763	5	3.7	270	1	YNF0_YEAST	P53952 saccharomyc	836	5	3.7	289	1	SGCD_MESAU	P97281 mesocricetu

837	5	3.7	289	1	SGCD_MOUSE	P82347	mus musculus	910	5	3.7	308	1	Y040_BP14	P39254	bacterioph
838	5	3.7	289	1	V3IK_TOBSV	P03537	tobacco str	911	5	3.7	308	1	Y222_MYCPN	P75466	mycoplasma
839	5	3.7	289	1	Y097_MYCTO	Q10893	mycobacteri	912	5	3.7	309	1	IFRH_MAIZE	P52500	zea mays (m
840	5	3.7	290	1	ARY1_CHICK	P13913	gallus gall	913	5	3.7	309	1	IK11_YEAST	P38874	saccharomyc
841	5	3.7	290	1	BIPI_TOBAC	003681	nicotiana t	914	5	3.7	309	1	MAT1_HUMAN	P51948	homo sapien
842	5	3.7	290	1	BLO2_KLEOX	P23954	klebsiella	915	5	3.7	309	1	MAT1_MOUSE	P51951	xenopus lae
843	5	3.7	290	1	CHIT_SACER	P14529	saccharopol	916	5	3.7	309	1	SCA1_XENLA	P53399	trichomonas
844	5	3.7	290	1	EPMO_RAT	P50279	rattus norv	917	5	3.7	309	1	SCA3_TRIVA	P53401	trichomonas
845	5	3.7	290	1	HIS1_SULSO	033771	sulfolobus	918	5	3.7	310	1	DP3B_MICLU	P21174	micrococcus
846	5	3.7	290	1	KDSA_PEA	050044	pisum sativ	919	5	3.7	310	1	YCYJ_ECOLI	P76049	escherichia
847	5	3.7	290	1	PARB_HELPJ	092K75	helicobacte	920	5	3.7	310	1	E2B1_MERTH	Q27900	methanobact
848	5	3.7	290	1	PARB_HELPJ	025758	helicobacte	921	5	3.7	311	1	PMG2_YEAST	Q12008	saccharomyc
849	5	3.7	290	1	RPOD_AERPE	Q9YB53	aeropyrum p	922	5	3.7	311	1	MDH_ECOLI	P06994	escherichia
850	5	3.7	290	1	SGCD_HUMAN	Q92629	homo sapien	923	5	3.7	312	1	MDH_ECOLI	P06994	escherichia
851	5	3.7	291	1	BLO1_KLEOX	P22391	klebsiella	924	5	3.7	312	1	OGG1_SALTY	P25077	salmonella
852	5	3.7	291	1	MLER_LACIA	P16400	lactococcus	925	5	3.7	312	1	TRXB_CHLMU	Q27337	methanobact
853	5	3.7	292	1	BIPI_TOBAC	003682	nicotiana t	926	5	3.7	312	1	TRXB_CHLTR	Q27337	methanobact
854	5	3.7	292	1	DAPA_ECO57	P58206	escherichia	927	5	3.7	312	1	YTF1_ECOS7	P58823	escherichia
855	5	3.7	292	1	DAPA_ECOLI	P05640	escherichia	928	5	3.7	312	1	YTF1_ECOS7	P58823	escherichia
856	5	3.7	292	1	GLMT_PIG	Q29555	sus scrofa	929	5	3.7	312	1	IF2A_HUMAN	P51598	homo sapien
857	5	3.7	292	1	GLMT_RABIT	Q29513	oryctolagus	930	5	3.7	314	1	IF2A_RAT	P51598	rattus norv
858	5	3.7	292	1	GLMT_RAT	P13255	rattus norv	931	5	3.7	314	1	REP1_STRAU	P14490	staphylococ
859	5	3.7	292	1	SUCD_RICPR	008371	rickettsia	932	5	3.7	314	1	REP1_STRAU	Q04013	saccharomyc
860	5	3.7	292	1	YK14_YEAST	Q02202	saccharomyc	933	5	3.7	315	1	ATPG_STYPI	Q05384	synechococ
861	5	3.7	293	1	BIPI_TOBAC	Q03686	nicotiana t	934	5	3.7	315	1	CALU_HUMAN	Q43852	homo sapien
862	5	3.7	293	1	YD11_MERJA	Q58707	methanococ	935	5	3.7	315	1	CALU_MOUSE	Q35887	mus musculu
863	5	3.7	293	1	YNP7_CABEL	P34560	caenorhabdi	936	5	3.7	315	1	FHUD_BACSU	P37800	bacillus su
864	5	3.7	294	1	DAPA_CAUCR	Q9A900	caulobacter	937	5	3.7	315	1	VN35_ROT11	Q03344	turkey rota
865	5	3.7	294	1	FBRL_TERTH	Q27200	tetrahymena	938	5	3.7	315	1	APBE_CHLMU	Q9PK27	chlamydia t
866	5	3.7	294	1	GLMT_HUMAN	Q14749	homo sapien	939	5	3.7	316	1	APBE_CHLTR	Q84080	chlamydia t
867	5	3.7	294	1	MP53_PHAQO	P56166	phalaris aq	940	5	3.7	316	1	BLA3_BACCE	P06548	bacillus ce
868	5	3.7	295	1	GLUB_CORGL	P48242	corynebacte	941	5	3.7	316	1	GBLP_NEUCR	Q01369	neutrospora
869	5	3.7	295	1	LPXC_HELPJ	Q9ZM50	helicobacte	942	5	3.7	316	1	PTA_RHIME	Q9X448	rhizobium m
870	5	3.7	295	1	Y234_MERJA	Q25632	helicobacte	943	5	3.7	316	1	TAL_VIBCH	Q9K186	vibrio chol
871	5	3.7	295	1	Y234_MERJA	Q60289	methanococ	944	5	3.7	316	1	YCZ0_YEAST	P25606	saccharomyc
872	5	3.7	296	1	SMDF_HUMAN	Q15491	homo sapien	945	5	3.7	316	1	YX12_CABEL	Q11123	caenorhabdi
873	5	3.7	297	1	LE33_CABEL	Q09581	caenorhabdi	946	5	3.7	316	1	APE_MACRA	P10517	macaca fasc
874	5	3.7	297	1	PERR_ECOLI	Q57083	escherichia	947	5	3.7	317	1	APE_PAPAN	P05770	papio anubi
875	5	3.7	297	1	XLVA_BACSU	P39800	bacillus su	948	5	3.7	317	1	PIX2_HUMAN	Q99697	homo sapien
876	5	3.7	297	1	YD67_MERJA	Q58762	methanococ	949	5	3.7	317	1	PIX2_MOUSE	P17408	mus musculu
877	5	3.7	297	1	YIM4_YEAST	P40471	saccharomyc	950	5	3.7	317	1	TLM_MOUSE	P17408	mus musculu
878	5	3.7	298	1	LAH1_SCHPO	P87056	schizosacch	951	5	3.7	317	1	YC05_YEAST	Q66883	aquifex aeo
879	5	3.7	298	1	YNO5_YEAST	P53843	saccharomyc	952	5	3.7	317	1	ACCA_ECOLI	P25616	saccharomyc
880	5	3.7	298	1	YNT6_YEAST	P40167	saccharomyc	953	5	3.7	317	1	ATPS_YEAST	P30867	escherichia
881	5	3.7	300	1	BLAB_PROVD	P52664	proteus vul	954	5	3.7	318	1	MYOD_HUMAN	P10085	mus musculu
882	5	3.7	300	1	YE03_YEAST	Q69782	rhizobium m	955	5	3.7	318	1	MYOD_MOUSE	Q02346	rattus norv
883	5	3.7	300	1	KIRP_BP14	P06855	bacterioph	956	5	3.7	318	1	MYOD_HUMAN	P15172	homo sapien
884	5	3.7	301	1	YGBJ_HAETN	P44979	haemophilus	957	5	3.7	318	1	MYOD_HUMAN	P15172	homo sapien
885	5	3.7	301	1	YGBJ_HAETN	Q9VEPI	dirosophila	958	5	3.7	318	1	MYOD_HUMAN	P15172	homo sapien
886	5	3.7	302	1	RRM1_DROME	P46805	thermotoga	959	5	3.7	318	1	MYOD_HUMAN	P15172	homo sapien
887	5	3.7	303	1	DDL_THEMA	Q33844	treponema p	960	5	3.7	318	1	MYOD_HUMAN	P15172	homo sapien
888	5	3.7	303	1	ROC_TREPA	P07910	homo sapien	961	5	3.7	318	1	MYOD_HUMAN	P15172	homo sapien
889	5	3.7	304	1	BLAI_AERHY	Q29005	aeromonas h	962	5	3.7	319	1	MYOD_HUMAN	P15172	homo sapien
890	5	3.7	304	1	IGIR_PIG	Q29005	sus scrofa	963	5	3.7	319	1	MYOD_HUMAN	P15172	homo sapien
891	5	3.7	304	1	INO2_YEAST	P26796	saccharomyc	964	5	3.7	319	1	MYOD_HUMAN	P15172	homo sapien
892	5	3.7	305	1	ILIVE_SYNY3	P54691	synechocyst	965	5	3.7	319	1	MYOD_HUMAN	P15172	homo sapien
893	5	3.7	305	1	MIAA_AQUAE	Q67162	aquifex aeo	966	5	3.7	319	1	MYOD_HUMAN	P15172	homo sapien
894	5	3.7	305	1	PPX1_ARATH	P48528	arabidopsis	967	5	3.7	319	1	MYOD_HUMAN	P15172	homo sapien
895	5	3.7	305	1	PPX2_ARATH	P48528	arabidopsis	968	5	3.7	319	1	MYOD_HUMAN	P15172	homo sapien
896	5	3.7	305	1	YHCH_BACSU	P45492	bacillus su	969	5	3.7	319	1	MYOD_HUMAN	P15172	homo sapien
897	5	3.7	306	1	BLAC_BACAM	Q44674	bacillus am	970	5	3.7	319	1	MYOD_HUMAN	P15172	homo sapien
898	5	3.7	306	1	BLAC_BACAM	P39824	bacillus su	971	5	3.7	319	1	MYOD_HUMAN	P15172	homo sapien
899	5	3.7	307	1	BLAC_BACLI	P00808	bacillus li	972	5	3.7	319	1	MYOD_HUMAN	P15172	homo sapien
900	5	3.7	307	1	COPE_CRIGR	Q60405	criceulius	973	5	3.7	319	1	MYOD_HUMAN	P15172	homo sapien
901	5	3.7	307	1	KHSE_DEIRA	Q98899	delinococcus	974	5	3.7	319	1	MYOD_HUMAN	P15172	homo sapien
902	5	3.7	307	1	MMP1_MYCAV	P46841	mycobacteri	975	5	3.7	319	1	MYOD_HUMAN	P15172	homo sapien
903	5	3.7	307	1	MMP1_MYCLE	P46841	mycobacteri	976	5	3.7	319	1	MYOD_HUMAN	P15172	homo sapien
904	5	3.7	307	1	Q0X2_ACAC	P70362	mus musculu	977	5	3.7	319	1	MYOD_HUMAN	P15172	homo sapien
905	5	3.7	307	1	UFDI_MOUSE	Q91165	bacterioph	978	5	3.7	319	1	MYOD_HUMAN	P15172	homo sapien
906	5	3.7	307	1	VP53_BPAPS	Q07698	aeromonas s	979	5	3.7	319	1	MYOD_HUMAN	P15172	homo sapien
907	5	3.7	308	1	ABCA_AERSA	Q58866	methanococ	980	5	3.7	319	1	MYOD_HUMAN	P15172	homo sapien
908	5	3.7	308	1	METE_MERJA	Q58866	methanococ	981	5	3.7	319	1	MYOD_HUMAN	P15172	homo sapien
909	5	3.7	308	1	METE_MERJA	Q58866	methanococ	982	5	3.7	319	1	MYOD_HUMAN	P15172	homo sapien

```

983 5 3.7 322 1 SLPI_DROME P32030 drosophila
984 5 3.7 322 1 TRXB_SCHPO Q92375 schizosacch
985 5 3.7 323 1 CRO_RANCA P17264 rana catesb
986 5 3.7 323 1 CRO_RANTE P02532 rana tempor
987 5 3.7 323 1 G731_HUMAN P09758 homo sapien
988 5 3.7 323 1 PIMI_XENLA Q91822 xenopus lae
989 5 3.7 323 1 PREA_CYACA Q91811 cyandidium c
990 5 3.7 324 1 E2B2_PYPAB Q9V281 pyrococcus
991 5 3.7 324 1 I12B_PIG Q28338 sus scrofa
992 5 3.7 324 1 Y700_METJA Q58111 methanococ
993 5 3.7 324 1 YB1_HUMAN P16991 homo sapien
994 5 3.7 324 1 YF02_METJA Q58897 methanococ
995 5 3.7 326 1 UNO3_HUMAN P22674 homo sapien
996 5 3.7 326 1 YOB3_HUMAN P45950 bacillus su
997 5 3.7 327 1 EBGR_ECOLI P06846 escherichia
998 5 3.7 327 1 FBRL_GIALA Q24957 giardia lam
999 5 3.7 327 1 FBRL_YEAST P15646 saccharomyc
1000 5 3.7 327 1 I12B_BOVIN P46282 bos taurus

```

## ALIGNMENTS

## RESULT 1

SAK\_STAU STANDARD: PRT: 163 AA.

AC P00802: 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE STAPHYLOKINASE PRECURSOR (NEUTRAL PROTEINASE) (PROTEASE III).  
 GN SAK.

OS Staphylococcus aureus.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Staphylococcus.  
 OX NCBI\_TaxID=1280;  
 [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=84069795; PubMed=6359061;  
 RA Sako T., Tsuchida N.;

RT "Nucleotide sequence of the staphylokinase gene from Staphylococcus  
 RT aureus.";  
 RL Nucleic Acids Res. 11:7679-7693(1983).  
 [2]

RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 43-163.  
 RX MEDLINE=97290447; PubMed=9145104;  
 RA Rabijs A., de Bondt H.L., de Rantier C.;

RT "Three-dimensional structure of staphylokinase, a plasminogen  
 RT activator with therapeutic potential.";  
 RL Nat. Struct. Biol. 4:357-360(1997).  
 [3]

RP STRUCTURE BY NMR OF 28-163  
 RX MEDLINE=96367505; PubMed=9692953;  
 RA Ohlenschlaeger O., Ramchandran R., Guehrs K.H., Schloft B.,  
 RA Brown L.R.;

RT "Nuclear magnetic resonance solution structure of the plasminogen-  
 RT activator protein staphylokinase.";  
 RL Biochemistry 37:10635-10642(1998).  
 -1- FUNCTION: THIS EXTRACELLULAR PROTEIN IS ONE OF THE PLASMINOGEN

CC ACTIVATORS THAT CONVERTS PLASMINOGEN INTO PLASMIN. IT REQUIRES  
 CC CALCIUM ION FOR STABILIZATION.  
 CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC

DR EMBL; X00127; CAA24957.1; -;  
 DR PIR; A00995; PRSAK.  
 DR PDB; 2SAK; 25-FEB-98.

DR PDB; 1SSN; 02-DEC-98.  
 KW Hydroxylase; Calcium; Plasminogen activation; Signal; 3D-structure.  
 FT SIGNAL 1  
 FT CHAIN 28 163  
 SQ SEQUENCE 163 AA; 18490 MW; E5609FF50ADE141 CRC64;

Query Match 75.0%; Score 102; DB 1; Length 163;  
 Best Local Similarity 100.0%; Pred. No. 1,4e-97;  
 Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 35 KGNELISPHYVEPIPGTTLKEKIEYVEMALDTAVKERRVELDPSAKIEVYDK 94
DB 62 KGNELISPHYVEPIPGTTLKEKIEYVEMALDTAVKERRVELDPSAKIEVYDK 121
QY 95 NKKKEETKSPITEKGVEVPDLSEHIKNPGLITKVIEKK 136
DB 122 NKKKEETKSPITEKGVEVPDLSEHIKNPGLITKVIEKK 163

```

## RESULT 2

SAK\_BPP42 STANDARD: PRT: 163 AA.

AC P15240: 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE STAPHYLOKINASE PRECURSOR (NEUTRAL PROTEINASE) (PROTEASE III).  
 GN SAK.

OS Bacteriophage P42D.  
 OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;  
 OC Lambda phage group.  
 OX NCBI\_TaxID=10715;  
 [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=88121731; PubMed=3123893;  
 RA Behnke D., Gerlach D.;

RT "Cloning and expression in Escherichia coli, Bacillus subtilis, and  
 RT Streptococcus sanguis of a gene for staphylokinase -- a bacterial  
 RT plasminogen activator.";  
 RL Mol. Gen. Genet. 210:528-534(1987).  
 -1- FUNCTION: THIS EXTRACELLULAR PROTEIN IS ONE OF THE PLASMINOGEN

CC ACTIVATORS THAT CONVERTS PLASMINOGEN INTO PLASMIN. IT REQUIRES  
 CC CALCIUM ION FOR STABILIZATION.  
 CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC

DR EMBL; X06603; CAA29822.1; -;  
 DR EMBL; M57455; AA98206.1; -;  
 DR EMBL; A17537; CAA01341.1; -;  
 DR PIR; S02330; S02330.  
 DR HSSP; P00802; 2SAK.

KW Hydroxylase; Calcium; Plasminogen activation; Signal.  
 FT SIGNAL 1 27  
 FT CHAIN 28 163  
 FT SEQUENCE 163 AA; 18608 MW; AA267A4FF75D36EC CRC64;

Query Match 68.4%; Score 93; DB 1; Length 163;  
 Best Local Similarity 100.0%; Pred. No. 2,6e-88;  
 Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 44 YVEPIPGTTLKEKIEYVEMALDTAVKERRVELDPSAKIEVYDKNKKKEETKS 103
DB 71 YVEPIPGTTLKEKIEYVEMALDTAVKERRVELDPSAKIEVYDKNKKKEETKS 130
QY 104 FPITEKGVEVPDLSEHIKNPGLITKVIEKK 136

```

DB 131 PPTKEGVVDDLSEHINKFNLITKVIIEKK 163

|||||

RESULT 3

TRPA\_HALN1 STANDARD; PRT; 204 AA.

AC Q9H8B9;

DT 20-AUG-2001 (Rel. 40, Created)

DT 20-AUG-2001 (Rel. 40, Last sequence update)

DE TRYPHOPHAN SYNTHASE ALPHA CHAIN (EC 4.2.1.20).

GN TRPA OR VNG03086.

OS Halobacterium sp. (strain NRC-1).

OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae; Halobacterium.

OC NCBI\_Taxid=64091;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20504483; PubMed=11016950;

RA Ng W.V., Kennedy S.P., Mahatras G.G., Bergquist B., Pan M., Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shroana J., Swartzell S., Weir D., Hall J., Dahl T.A., Weir R., Goo Y.A., Leitbauer B., Keller K., Cruz R., Danson M.J., Hough D.W., Madocks D.G., Jablonki P.E., Krebs M.P., Angevine C.M., Dale H., Idenberger T.A., Peck R.F., Pohlischer M., Spudich J.L., Jung K.-H., Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D., Rahardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;

RT "genome sequence of Halobacterium species NRC-1";

RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).

CC -1- FUNCTION: THE ALPHA SUBUNIT IS RESPONSIBLE FOR THE ALDOL CLEAVAGE OF INDOLGLYCEROL PHOSPHATE TO INDOL AND GLYCERALDEHYDE 3-PHOSPHATE.

CC PROSHATE.

CC -1- CATALYTIC ACTIVITY: L-SERINE + 1-(INDOL-3-YL)GLYCEROL 3-PHOSPHATE = L-TRYPTOPHAN + GLYCERALDEHYDE 3-PHOSPHATE + H(2)O.

CC -1- PATHWAY: LAST (FIFTH) STEP IN BIOSYNTHESIS OF TRYPTOPHAN.

CC -1- SUBUNIT: TRIMER OF TWO ALPHA AND TWO BETA CHAINS (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE TRPA FAMILY.

CC -----

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: AE004991; AAC18888.1; -

DR InterPro: IPR003009; FMN\_enzyme.

DR InterPro: IPR002028; TRP\_synthase\_alpha.

DR Pfam: PF00290; trp\_synta.1.

DR ProDom: PD001535; TRP\_synthase\_alpha.1.

DR PROSITE: PS00167; TRP\_SYNTHASE\_ALPHA; FALSE NEG.

KW Tryptophan biosynthesis; lyase; Complete proteome.

SO SEQUENCE 204 AA; 21172 MW; 7723CC4483A5ED0C CRC64;

Query Match 5.1%; Score 7; DB 1; Length 204; Best Local Similarity 100.0%; Pred. No. 7.5; Mismatches 0; Indels 0; Gaps 0; Matches 7; Conservative 0;

QY 110 GFVVDL 116

Db 51 GFVVDL 57

RESULT 4

CTRA\_CAUCR STANDARD; PRT; 231 AA.

ID CTRA\_CAUCR

AC 045994;

DT 20-AUG-2001 (Rel. 40, Created)

DT 20-AUG-2001 (Rel. 40, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE CELL CYCLE TRANSCRIPTIONAL REGULATOR CTRA (RESPONSE REGULATOR SOKA).

GN CTRA OR SOKA OR CC3035.

OS Caulobacter crescentus.

OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group; Caulobacter.

OC NCBI\_Taxid=69394;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-CB15N / NA1000;

RX MEDLINE=96140642; PubMed=8548829;

RA Oton K.C., Marcynski G.T., Shapiro L.;

RT "cell cycle control by an essential bacterial two-component signal transduction protein";

RL Cell 84:83-93(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 19089 / CB15;

RX MEDLINE=21173698; PubMed=11259647;

RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E., Eisen J., Heidelberg J.F., Alley M.K.R., Ohta N., Maddock J.R., Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B., Deboy J.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H., Kolony J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K., Uterback T., Tran K., Wolf A., Yamathayan J., Ermolaeva M., White O., Salzberg S.L., Shapiro L., Venter J.C., Fraser C.M.;

RT "Complete genome sequence of Caulobacter crescentus";

RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).

RN [3]

RP SEQUENCE OF 159-231 FROM N.A.

RC STRAIN-ATCC 19089 / CB15;

RX MEDLINE=98132609; PubMed=9465034;

RA Wu J., Ohta N., Newton A.;

RT "An essential, multicomponent signal transduction pathway required for cell cycle regulation in Caulobacter";

RL Proc. Natl. Acad. Sci. U.S.A. 95:1443-1448(1998).

CC -1- FUNCTION: FORMS PART OF A TWO-COMPONENT REGULATORY SYSTEM CTRA/CCRA THAT CONTROLS MULTIPLE EVENTS IN THE CELL CYCLE, INCLUDING CELL DIVISION, STALK SYNTHESIS, AND CELL CYCLE-SPECIFIC TRANSCRIPTION. BINDS TO A GROUP OF CELL CYCLE-REGULATED PROMOTERS CRITICAL FOR DNA REPLICATION, DNA METHYLATION, AND CLASS II FLAGELLAR BIOGENESIS.

CC -1- PMM: PHOSPHORYLATED BY CCKA.

CC -1- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS.

CC -----

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: U39559; AAA93080.1; -

DR EMBL: AE005966; AAK24997.1; -

DR EMBL: AF021339; AAC05479.1; -

DR TIGR: CC3035; -

DR InterPro: IPR001789; Response\_reg.

DR InterPro: IPR001867; Trans\_reg\_C.

DR Pfam: PF00072; response\_reg.1.

DR Pfam: PF00486; trans\_reg\_C.1.

DR SMART: SM00448; REC.1.

KW Sensory transduction; Phosphorylation; Transcription regulation; DNA binding; Complete proteome.

FT MOD\_RES 51

SO SEQUENCE 231 AA; 25796 MW; BB812AF6EDADBF8 CRC64;

Query Match 5.1%; Score 7; DB 1; Length 231; Best Local Similarity 100.0%; Pred. No. 8.4; Mismatches 0; Indels 0; Gaps 0; Matches 7; Conservative 0;

OY 52 GTTTC 58  
GN |||||  
DB 166 GTTTC 172

## RESULT 5

TRPA\_HALVO STANDARD: PRT; 277 AA.  
AC P18284;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE TRYPTOPHAN SYNTHASE ALPHA CHAIN (EC 4.2.1.20).  
GN TRPA.  
OS Halobacterium volcanii (Halobacter volcanii).  
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae; Haloferax.  
OX NCBI\_TaxID=2246;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90370836; PubMed=2118654;  
RA Lam W.L., Cohen A., Tsoulfas D., Doolittle W.F.;  
RT "Genes for tryptophan biosynthesis in the archaeobacterium Haloferax  
RT volcanii.";  
RT Proc. Natl. Acad. Sci. U.S.A. 87:6614-6618(1990).  
CC -1- FUNCTION: THE ALPHA SUBUNIT IS RESPONSIBLE FOR THE ALDOL CLEAVAGE  
CC OF INDOLBUTYRYL PHOSPHATE TO INDOL AND GLYCERALDEHYDE 3-  
CC PHOSPHATE.  
CC -1- CATALYTIC ACTIVITY: L-SERINE + L-(INDOL-3-YL)GLYCEROL 3-PHOSPHATE  
CC = L-TRYPTOPHAN + GLYCERALDEHYDE 3-PHOSPHATE + H(2)O.  
CC -1- PATHWAY: LAST (FIFTH) STEP IN BIOSYNTHESIS OF TRYPTOPHAN.  
CC -1- SUBUNIT: TETRAMER OF TWO ALPHA AND TWO BETA CHAINS (BY  
CC SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE TRPA FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M36177; AAA72864.1; -.  
DR PIR: A36044; A36044.  
DR HSSP: P00929; 1BKS.  
DR InterPro: IPR003009; FMN\_enzyme.  
DR InterPro: IPR002028; TRP\_synthase\_alpha.  
DR Pfam: PF00290; trp\_synth. 1.  
DR PROSITE: PS001535; TRP\_synthase\_alpha; 1.  
DR PROSITE: PS00167; TRP\_synthase\_alpha; 1.  
KW Tryptophan biosynthesis; Lyase.  
FT ACT\_SITE 54 54 INVOLVED IN PROTON DONATION/ACCEPTATION  
FT BINDING 205 208 SUBSTRATE.  
FT DURING CATALYSIS.  
FT SEQUENCE 277 AA; 29720 MW; 22F68CE6826B0E61 CRC64;  
SQ

Query Match 5.1%; Score 7; DB 1; Length 277;  
Best Local Similarity 100.0%; Pred. No. 9.8;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 110 GFVVPDL 116  
DB 120 GFVVPDL 126

## RESULT 6

TRXB\_NEUCR STANDARD: PRT; 334 AA.  
AC P51978;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE THIOREDOXIN REDUCTASE (EC 1.6.4.5).  
GN CYS-9.  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariales; Sordariaceae; Neurospora.  
OX NCBI\_TaxID=5141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ACR-2;  
RA Onai K., Takayanagi K., Nakashima H.;  
RT Submitted (Apr-1995) to the EMBL/Genbank/DDJ databases.  
CC -1- CATALYTIC ACTIVITY: NADPH + OXIDIZED THIOREDOXIN = NADP(+) +  
CC REDUCED THIOREDOXIN.  
CC -1- COFACTOR: FAD.  
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).  
CC -1- MISCELLANEOUS: THE ACTIVE SITE IS A REDOX-ACTIVE DISULFIDE BOND.  
CC -1- SIMILARITY: BELONGS TO THE PYRIDINE NUCLEOTIDE-DISULFIDE  
CC OXIDOREDUCTASES CLASS-II.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: D45049; BAA08090.1; -.  
DR HSSP: Q39243; 1VDC.  
DR InterPro: IPR001327; FAD\_Pyr\_redox.  
DR InterPro: IPR00103; Pyridine\_redox.2.  
DR Pfam: PF00070; Pyr\_redox; 1.  
DR PRINTS: PR00368; FADPNR.  
DR PRINTS: PR00469; PNDPDTSEII.  
DR PROSITE: PS00573; PYRIDINE\_REDox\_2; 1.  
KW Redox-active center; Oxidoreductase; NADP; Flavoprotein; FAD.  
FT NP\_BIND 7 24 FAD (ADP PART) (PROBABLE).  
FT DISULFID 140 143 REDOX-ACTIVE (BY SIMILARITY).  
FT NP\_BIND 155 168 NAD(P) (BY SIMILARITY).  
FT NP\_BIND 277 309 FAD (FLAVIN PART) (BY SIMILARITY).  
FT SEQUENCE 334 AA; 35882 MW; 4D9E9861E479EB4 CRC64;  
SQ

Query Match 5.1%; Score 7; DB 1; Length 334;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 50 KPCTILT 56  
DB 271 KPCTILT 277

## RESULT 7

PBPL\_YEAST STANDARD: PRT; 722 AA.  
AC P53297;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE PBPL-BINDING PROTEIN 1.  
GN PBPL OR MRS16 OR YGR178C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hedling U., Holmann B., Delius H.;  
RT Submitted (May-1996) to the EMBL/Genbank/DDJ databases.  
RN [12]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DBY747;



RA Mecklenbrauker I.;  
RT "Sequencing and characterization of a suppressor of the pet<sup>-</sup> phenotype  
in a Saccharomyces cerevisiae strain without mitochondrial group II  
introns";  
RL Theiss (1996), Vienna Biocentre, Austria.  
RN [3]  
RP CHARACTERIZATION.  
RX MEDLINE=99038243; PubMed=9819425;  
RA Mangus D.A., Amrani N., Jacobson A.;  
RT "Pbplp, a factor interacting with Saccharomyces cerevisiae poly(A)-  
binding protein, regulates polyadenylation.";  
RL Mol. Cell. Biol. 18:7383-7396(1998)  
CC -1- FUNCTION: APPEARS TO PROMOTE PROPER POLYADENYLATION. IN THE  
CC ABSENCE OF PBPLP, THE 3'TERMINI OF PRE-MRNAs ARE PROPERLY CLEAVED  
CC BUT LACK FULL-LENGTH POLY(A) TAILS. MAY ACT TO REPRESS THE ABILITY  
CC OF PAB1 TO NEGATIVELY REGULATE POLYADENYLATION.  
CC -1- SUBUNIT: INTERACTS WITH PAB1.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR AND CYTOPLASMIC.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Z72963; CA97204.1; -  
DR EMBL; U46931; AAB94294.1; -  
DR SGD; S0003410; PBPL.  
KM NUCLEAR protein.  
SQ SEQUENCE 722 AA; 78781 MW; 92005F3A2346193E CRC64;  
  
Query Match 5.1%; Score 7; DB 1; Length 722;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 10 KKGDAS 16  
|||  
Db 514 KKGDAS 520  
  
RESULT 8  
ID IF2\_HELPY STANDARD; PRT; 944 AA.  
AC P55972;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE TRANSLATION INITIATION FACTOR IF-2.  
GN INFB OR HP1048.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
CC Helicobacter.  
CC NCBL\_TaxID=210;  
OX [1]  
RN  
RP SEQUENCE FROM N.A.  
RX STRAIN=26695 / ATCC 700392;  
RX MEDLINE=97394467; PubMed=9252185;  
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,  
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,  
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,  
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,  
RA McEwen K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,  
RA Berg D.E., Gocayne J.D., Uterback T.R., Peterson J.D., Kelley J.M.,  
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,  
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,  
RA Venter J.C.;  
RT "The complete genome sequence of the gastric pathogen Helicobacter  
RT pylori.";  
RL Nature 388:539-547(1997).  
CC -1- FUNCTION: IF-2, ONE OF THE ESSENTIAL COMPONENTS FOR THE INITIATION

CC OF PROTEIN SYNTHESIS IN VITRO, PROTECTS FORMYL METHIONYL-TRNA FROM  
CC SPONTANEOUS HYDROLYSIS AND PROMOTES ITS BINDING TO THE 30S  
CC RIBOSOMAL SUBUNITS. IT IS ALSO INVOLVED IN THE HYDROLYSIS OF GTP  
CC DURING THE FORMATION OF THE 70S RIBOSOMAL COMPLEX (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- SIMILARITY: BELONGS TO THE IF-2 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AE000612; AAD08093.1; -  
DR HSSP; P02990; IEFU.  
DR TIGR; HP1048; -  
DR InterPro; IPR000795; GTP\_EFTU.  
DR InterPro; IPR000178; IF2.  
DR Pfam; PF00009; GTP\_EFTU; 1.  
DR Pfam; PF02131; IF2; 1.  
DR ProDom; PD186100; IF2; 1.  
DR PROSITE; PS01176; IF2; 1.  
KM Initiation factor; Protein biosynthesis; GTP-binding;  
KW Complete proteome.  
FT DOMAIN 446 594 G-DOMAIN.  
FT NP\_BIND 452 459 GTP (BY SIMILARITY).  
FT NP\_BIND 498 502 GTP (BY SIMILARITY).  
FT NP\_BIND 552 555 GTP (BY SIMILARITY).  
SQ SEQUENCE 944 AA; 105194 MW; EC90404D9C3B851D CRC64;  
  
Query Match 5.1%; Score 7; DB 1; Length 944;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 97 KKEFTKS 103  
|||  
Db 96 KKEFTKS 102  
  
RESULT 9  
ID POL\_HV2CA STANDARD; PRT; 1034 AA.  
AC P24107;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE POL POLYPEPTIDE [CONTAINS: PROTEASE (RETROPEPSIN) (EC 3.4.23.16);  
DE REVERSE TRANSCRIPTASE (EC 2.7.7.49); RIBONUCLEASE H (EC 3.1.26.4)].  
GN POL.  
OS Human immunodeficiency virus type 2 (isolate CAM2) (HIV-2).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
CC NCBL\_TaxID=11715;  
OX [1]  
RN  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91170959; PubMed=2005437;  
RT "Nucleotide sequence of a Guinea-Bissau-derived human  
RT immunodeficiency virus type 2 proviral clone (HIV-2CAM2).";  
RL J. Gen. Virol. 72:721-724(1991).  
CC -1- P1M: CLEAVAGE SITES THAT YIELD THE MATURE PROTEINS REMAIN TO BE  
CC DETERMINED.  
CC -1- MISCELLANEOUS: READTHROUGH OF THE TERMINATOR TGA MAY OCCUR BETWEEN  
CC 564-ILE AND 565-GLY.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO  
CC KNOWN AS THE RETROPEPSIN FAMILY.  
CC PIR; B38475; GNJCJA.  
DR HSSP; P04584; IJLD.  
DR MEROPS; A02.002; -  
DR InterPro; IPR001995; Asp\_prot\_retrov.  
DR InterPro; IPR001969; Asp\_protease.

DR InterPro: IPR001037; Integrase\_C.  
 DR InterPro: IPR003308; Integrase\_zn.  
 DR InterPro: IPR002156; RNaseH.  
 DR InterPro: IPR000477; RVTse.  
 DR InterPro: IPR001584; Rve.  
 DR Pfam: PF00552; Integrase; 1.  
 DR Pfam: PF02022; Integrase\_zn; 1.  
 DR Pfam: PF00075; RNaseH; 1.  
 DR Pfam: PF00665; rve; 1.  
 DR Pfam: PF00077; rvp; 1.  
 DR Pfam: PF00078; rvt; 1.  
 DR PROSITE: PS00141; ASP\_PROTEASE; 1.  
 DR PROSITE: PS0175; ASP\_PROT\_RETROV; 1.  
 DR AIDS: Polypotein; Hydrolyase; Aspartyl protease; Endonuclease;  
 DR Nuclease; Transferase; RNA-directed DNA polymerase.  
 FT CHAIN 85  
 FT ACT SITE 109 109 BY SIMILARITY.  
 SO SEQUENCE 1034 AA; 117195 MW; 3514E566AA6D7C86 CRC64;

Query Match 5.1%; Score 7; DB 1; Length 1034;  
 Best Local Similarity 100.0%; Pred. No. 31;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 55 LTRKITE 61  
 DB 209 LTRKITE 215

RESULT 10  
 POL\_HV2KR STANDARD; PRT; 1035 AA.  
 ID 074120;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE POLYPROTEIN [CONTAINS: PROTEASE (RETROPEPSIN) (EC 3.4.23.16);  
 DE REVERSE TRANSCRIPTASE (EC 2.7.7.49); RIBONUCLEASE H (EC 3.1.26.4)].  
 GN POL.  
 OS Human immunodeficiency virus type 2 (isolate RH-2) (HIV-2).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentiviruses.  
 NCBI\_TaxID=73484;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kraus G.K., Talbot R., Leavitt M., Luznick L., Schmidt A.,  
 RA Badel P., Barz C., Morton W., Wong-Schaal F., Looney D.J.;  
 RL Submitted (Apr-1995) to the EMBL/Genbank/DBJ databases.  
 CC -1- PWM: CLEAVAGE SITES THAT YIELD THE MATURE PROTEINS REMAIN TO BE  
 DETERMINED.  
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO  
 KNOWN AS THE RETROPEPSIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U22047; AAA64576.1; -.  
 DR HSSP: P04584; IJLD.  
 DR InterPro: IPR001995; Asp\_prot\_retrov.  
 DR InterPro: IPR001969; Asp\_protease.  
 DR InterPro: IPR001037; Integrase\_C.  
 DR InterPro: IPR003308; Integrase\_zn.  
 DR InterPro: IPR002156; RNaseH.  
 DR InterPro: IPR000477; RVTse.  
 DR InterPro: IPR001584; Rve.  
 DR Pfam: PF00552; Integrase; 1.  
 DR Pfam: PF02022; Integrase\_zn; 1.  
 DR Pfam: PF00075; RNaseH; 1.  
 DR Pfam: PF00077; rvp; 1.  
 DR Pfam: PF00078; rvt; 1.

DR Pfam: PF00077; rvp; 1.  
 DR Pfam: PF00078; rvt; 1.  
 DR PROSITE: PS00141; ASP\_PROTEASE; 1.  
 DR PROSITE: PS0175; ASP\_PROT\_RETROV; 1.  
 DR AIDS: Polypotein; Hydrolyase; Aspartyl protease; Endonuclease;  
 DR Nuclease; Transferase; RNA-directed DNA polymerase.  
 FT CHAIN 85  
 FT ACT SITE 109 109 BY SIMILARITY.  
 SO SEQUENCE 1035 AA; 117632 MW; 696EDC6CAFBD06CF CRC64;

Query Match 5.1%; Score 7; DB 1; Length 1035;  
 Best Local Similarity 100.0%; Pred. No. 31;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 55 LTRKITE 61  
 DB 209 LTRKITE 215

RESULT 11  
 POL\_HV2NZ STANDARD; PRT; 1035 AA.  
 ID AC P05962; Q85571;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE POLYPROTEIN [CONTAINS: PROTEASE (RETROPEPSIN) (EC 3.4.23.16);  
 DE REVERSE TRANSCRIPTASE (EC 2.7.7.49); RIBONUCLEASE H (EC 3.1.26.4)].  
 GN POL.  
 OS Human immunodeficiency virus type 2 (isolate NIH-2) (HIV-2).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 NCBI\_TaxID=11719;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=88320359; PubMed=3261862;  
 RA Zagury J.F., Franchini G., Reitz M.S. Jr., Collalti E., Starck B.R.,  
 RA Hall L., Fargnoli K., Jagodzinski L.L., Guo H.-G., Laure F.,  
 RA Arya S.K., Josephs S.F., Zagury D., Wong-Schaal F., Gallo R.C.;  
 RT "Genetic variability between isolates of human immunodeficiency virus  
 RT (HIV) type 2 is comparable to the variability among HIV type 1,"  
 RL Proc. Natl. Acad. Sci. U.S.A. 85:5941-5945(1988).  
 CC -1- PWM: CLEAVAGE SITES THAT YIELD THE MATURE PROTEINS REMAIN TO BE  
 DETERMINED.  
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO  
 KNOWN AS THE RETROPEPSIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: J03654; AAB00755.1; ALT\_INT.  
 DR HSSP: P04584; IJLD.  
 DR HIV: J03654; POLS2NHZ.  
 DR MEROPS: A02.002; -.  
 DR InterPro: IPR001995; Asp\_prot\_retrov.  
 DR InterPro: IPR001969; Asp\_protease.  
 DR InterPro: IPR001037; Integrase\_C.  
 DR InterPro: IPR003308; Integrase\_zn.  
 DR InterPro: IPR002156; RNaseH.  
 DR InterPro: IPR000477; RVTse.  
 DR InterPro: IPR001584; Rve.  
 DR Pfam: PF00552; Integrase; 1.  
 DR Pfam: PF02022; Integrase\_zn; 1.  
 DR Pfam: PF00075; RNaseH; 1.  
 DR Pfam: PF00077; rvp; 1.  
 DR Pfam: PF00078; rvt; 1.  
 DR PROSITE: PS00141; ASP\_PROTEASE; FALSE\_NEG.



OY 55 LTKERIE 61  
Db 210 LTKERIE 216

RESULT 13  
POL\_HV2G1 STANDARD; PRT; 1049 AA.  
AC P18042;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE POL. POLYPROTEIN [CONTAINS: PROTEASE (RETROPEPSIN) (EC 3.4.23.16);  
REVERSE TRANSCRIPTASE (EC 2.7.7.49); RIBONUCLEASE H (EC 3.1.26.4)].  
GN POL.  
OS Human immunodeficiency virus type 2 (isolate Ghana-1) (HIV-2).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_Taxid=11717;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90112350; PubMed=2611042;  
RA Hasegawa A., Tsujimoto H., Maki N., Ishikawa K.I., Miura T.,  
RA Fukusawa M., Miki K., Hayami M.,  
RT "Genomic divergence of HIV-2 from Ghana."  
RL AIDS Res. Hum. Retroviruses 5:593-604(1989).  
CC -1- PTM: CLEAVAGE SITES THAT YIELD THE MATURE PROTEINS REMAIN TO BE  
CC DETERMINED.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO  
CC KNOWN AS THE RETROPEPSIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL, M30895; AAA43933.1; -.  
DR PIR: J50328; GNLJGG.  
DR HSSP: P04584; IJLD.  
DR HIV: M30895; POLS2GHL.  
DR MEROPS: A02\_002; -.  
DR InterPro: IPR001995; Asp\_prot\_retrov.  
DR InterPro: IPR001969; Asp\_protease.  
DR InterPro: IPR001037; Integrase\_C.  
DR InterPro: IPR003308; Integrase\_zn.  
DR InterPro: IPR002156; RNaseH.  
DR InterPro: IPR000477; RVTse.  
DR InterPro: IPR001584; Rve.  
DR Pfam: PF00552; Integrase\_1.  
DR Pfam: PF02022; Integrase\_zn; 1.  
DR Pfam: PF0075; RNaseH; 1.  
DR Pfam: PF00665; rve; 1.  
DR Pfam: PF00077; rvp; 1.  
DR Pfam: PF00078; rvt; 1.  
DR PROSITE: PS00141; ASP\_PROTEASE; 1.  
DR PROSITE: PS0175; ASP\_PROT\_RETROV; 1.  
DR AIDS; Polyprotein; Hydrolase; Aspartyl protease; Endonuclease;  
KW Nuclease; Transferase; RNA-directed DNA polymerase.  
FT CHAIN 98 207  
FT ACT\_SITE 123 123 BY SIMILARITY.  
SQ SEQUENCE 1049 AA; 119087 MW; D0123D490E899FC3 CRC64;

Query Match 5.1%; Score 7; DB 1; Length 1049;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 55 LTKERIE 61  
Db 223 LTKERIE 229

RESULT 14  
POL\_HV2ST STANDARD; PRT; 1055 AA.  
AC P20876;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE POL. POLYPROTEIN [CONTAINS: PROTEASE (RETROPEPSIN) (EC 3.4.23.16);  
REVERSE TRANSCRIPTASE (EC 2.7.7.49); RIBONUCLEASE H (EC 3.1.26.4)].  
GN POL.  
OS Human immunodeficiency virus type 2 (isolate ST) (HIV-2).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_Taxid=11721;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90112662; PubMed=2296086;  
RA Kumar P., Hui H., Kappes J.C., Hagarty B.S., Hoxie J.A., Arya S.K.,  
RA Shaw G.M., Hahn B.H.;  
RT "Molecular characterization of an attenuated human immunodeficiency  
RT virus type 2 isolate."  
RL J. Virol. 64:890-901(1990).  
CC -1- PTM: CLEAVAGE SITES THAT YIELD THE MATURE PROTEINS REMAIN TO BE  
CC DETERMINED.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO  
CC KNOWN AS THE RETROPEPSIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL, M31113; -; NOT\_ANNOTATED\_CDS.  
DR PIR: B33943; GNLJST.  
DR HSSP: P04584; IJLD.  
DR HIV: M31113; POLS2ST.  
DR MEROPS: A02\_002; -.  
DR InterPro: IPR001995; Asp\_prot\_retrov.  
DR InterPro: IPR001969; Asp\_protease.  
DR InterPro: IPR001037; Integrase\_C.  
DR InterPro: IPR003308; Integrase\_zn.  
DR InterPro: IPR002156; RNaseH.  
DR InterPro: IPR000477; RVTse.  
DR InterPro: IPR001584; Rve.  
DR Pfam: PF00552; Integrase\_1.  
DR Pfam: PF02022; Integrase\_zn; 1.  
DR Pfam: PF0075; RNaseH; 1.  
DR Pfam: PF00665; rve; 1.  
DR Pfam: PF00077; rvp; 1.  
DR Pfam: PF00078; rvt; 1.  
DR PROSITE: PS00141; ASP\_PROTEASE; 1.  
DR PROSITE: PS0175; ASP\_PROT\_RETROV; 1.  
DR AIDS; Polyprotein; Hydrolase; Aspartyl protease; Endonuclease;  
KW Nuclease; Transferase; RNA-directed DNA polymerase.  
FT CHAIN 105 203  
FT ACT\_SITE 129 129 BY SIMILARITY.  
SQ SEQUENCE 1055 AA; 119767 MW; A943310E567869B1 CRC64;

Query Match 5.1%; Score 7; DB 1; Length 1055;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 55 LTKERIE 61  
Db 229 LTKERIE 235

RESULT 15

POL\_HV2D1  
ID POL\_HV2D1 STANDARD; PRT: 1073 AA.  
AC P1757;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE POL POLYPROTEIN [CONTAINS: PROTEASE (RETROPEPSIN) (EC 3.4.23.16);  
REVERSE TRANSCRIPTASE (EC 2.7.7.49); RIBONUCLEASE H (EC 3.1.26.4)].  
GN POL.  
OS Human immunodeficiency virus type 2 (isolate D194) (HIV-2).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11713;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89184631; PubMed=2467304;  
RA Kuehnelt H., von Briesen H., Dietrich U., Adamski M., Mix D.,  
Blesert L., Kreutz R., Immelman A., Henco K., Melchsner C.,  
Andreesen R., Gelderblom H., Ruebsamen-Waigmann H.;  
RT "Molecular cloning of two west African human immunodeficiency virus  
type 2 isolates that replicate well in macrophages: a Gambian  
isolate, from a patient with neurologic acquired immunodeficiency  
syndrome, and a highly divergent Ghanaian isolate.";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:2383-2387(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91045094; PubMed=2235509;  
RA Kuehnelt H., Kreutz R., Ruebsamen-Waigmann H.;  
RT "Nucleotide sequence of HIV-2D194, an isolate from a Gambian case of  
'neuro-AIDS', which showed excellent growth in macrophages.";  
RL Nucleic Acids Res. 18:6142-6142(1990).  
CC -1- PTM: CLEAVAGE SITES THAT YIELD THE MATURE PROTEINS REMAIN TO BE  
DETERMINED.  
CC -1- MISCELLANEOUS: THIS ISOLATE IS FROM A GAMBIAN CASE OF  
'NEURO-AIDS'.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO  
KNOWN AS THE RETROPEPSIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: J04542; AAA76841.2; -;  
DR EMBL: X52223; -: NOT\_ANNOTATED\_CDS.  
DR PIR: S12153; S12153.  
DR HSSP: P04584; 1TLD.  
DR HIV: J04542; POL52D194.  
DR MEROPS: A02.002; -;  
DR InterPro: IPR001995; Asp\_prot\_retrov.  
DR InterPro: IPR001969; Asp\_protease.  
DR InterPro: IPR001037; Integrase\_C.  
DR InterPro: IPR003308; Integrase\_zn.  
DR InterPro: IPR002156; RNaseH.  
DR InterPro: IPR001584; RYse.  
DR InterPro: IPR000477; RYse.  
DR Pfam: PF00552; Integrase\_1.  
DR Pfam: PF02022; Integrase\_zn; 1.  
DR Pfam: PF00075; RNaseH; 1.  
DR Pfam: PF00665; rve; 1.  
DR Pfam: PF00077; rvp; 1.  
DR Pfam: PF00078; rvt; 1.  
DR PROSITE: PS00141; ASP\_PROTEASE; 1.  
DR PROSITE: PS50175; ASP\_PROT\_RETROV; 1.  
KW AIDS; Polypeptide; Hydroxylase; Aspartyl protease; Endonuclease;  
Nuclease; Transferase; RNA-directed DNA polymerase.  
FT CHAIN 123 221  
FT ACT SITE 147 147  
FT BY SIMILARITY.  
SQ SEQUENCE 1073 AA; 121071 MW; B3AA099483B4B3CF CRC64;

Query Match 5.18; Score 7; DB 1; Length 1073;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 LTKKIE 61  
DB 247 LTKKIE 253

Search completed: April 22, 2002, 10:49:02  
Job time: 197 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: April 22, 2002, 10:45:25 ; Search time 24.86 Seconds

(without alignments)  
800.203 Million cell updates/sec

Title: US-09-601-490-1

Perfect score: 136  
Sequence: 1 SSSFDKGGKKKGDSAYEEP.....SEHKPGFNLTQVIEKK 136Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 473505 seqs, 146272329 residues

Word size : 0

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

SPREMBL.17:\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.rodent:\*  
12: sp.virus:\*  
13: sp.vertibrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	136	100.0	163	2	099SU7
2	100	73.5	136	2	033929
3	97	71.3	163	2	09LC46
4	93	68.4	163	2	09AM04
5	8	5.9	245	2	09X260
6	7	5.1	56	1	027388
7	7	5.1	90	6	028835
8	7	5.1	92	12	086603
9	7	5.1	102	2	09K8J5
10	7	5.1	114	12	09JEH2
11	7	5.1	138	2	09KFP3
12	7	5.1	140	12	09OU35
13	7	5.1	154	12	09WU01
14	7	5.1	181	5	09NE98
15	7	5.1	185	2	09ACCO
16	7	5.1	194	2	084552
17	7	5.1	219	5	045975
18	7	5.1	221	12	09IVA8
19	7	5.1	224	12	098X05

20	7	5.1	224	12	098W5	098W5 human immun
21	7	5.1	232	2	09ZHS1	09ZHS1 Brucella ab
22	7	5.1	233	2	09FDP5	09FDP5 rhizobium m
23	7	5.1	236	12	090226	090226 turnip mosa
24	7	5.1	237	2	09RNH4	09RNH4 rhodobacter
25	7	5.1	250	12	0991V8	0991V8 human immun
26	7	5.1	263	10	09MSK7	09MSK7 glycine max
27	7	5.1	276	10	09SG92	09SG92 arbidops1s
28	7	5.1	288	12	088917	088917 turnip mosa
29	7	5.1	288	12	088599	088599 turnip mosa
30	7	5.1	288	12	P89215	P89215 turnip mosa
31	7	5.1	288	12	09IC16	09IC16 turnip mosa
32	7	5.1	288	12	09IC15	09IC15 turnip mosa
33	7	5.1	288	12	09IC14	09IC14 turnip mosa
34	7	5.1	288	12	09DKY5	09DKY5 turnip mosa
35	7	5.1	288	12	09DD00	09DD00 turnip mosa
36	7	5.1	288	12	099DC9	099DC9 turnip mosa
37	7	5.1	288	12	099PC8	099PC8 turnip mosa
38	7	5.1	298	12	09WA68	09WA68 turnip mosa
39	7	5.1	298	12	09WA67	09WA67 turnip mosa
40	7	5.1	298	12	09WA66	09WA66 turnip mosa
41	7	5.1	298	12	09WA64	09WA64 turnip mosa
42	7	5.1	298	12	09WA63	09WA63 turnip mosa
43	7	5.1	298	12	09WA62	09WA62 turnip mosa
44	7	5.1	298	12	09WA61	09WA61 turnip mosa
45	7	5.1	303	13	09IB66	09IB66 xenopus lae
46	7	5.1	311	2	09KET0	09KET0 bacillus ha
47	7	5.1	332	2	09XD02	09XD02 clostridium
48	7	5.1	341	12	09BWR7	09BWR7 human immun
49	7	5.1	343	1	09YDE0	09YDE0 aeropyrum p
50	7	5.1	343	12	09J5A2	09J5A2 fowlpox vir
51	7	5.1	349	12	09EAK0	09EAK0 human immun
52	7	5.1	351	12	09E3T6	09E3T6 human immun
53	7	5.1	351	12	09E3T5	09E3T5 human immun
54	7	5.1	351	12	09E3T4	09E3T4 human immun
55	7	5.1	351	12	09E3T3	09E3T3 human immun
56	7	5.1	351	12	09E3T2	09E3T2 human immun
57	7	5.1	351	12	09E3T1	09E3T1 human immun
58	7	5.1	353	12	09IGK4	09IGK4 human immun
59	7	5.1	361	12	099EM9	099EM9 turnip mosa
60	7	5.1	361	12	099EM8	099EM8 turnip mosa
61	7	5.1	361	12	099EM7	099EM7 turnip mosa
62	7	5.1	370	2	09KZV9	09KZV9 streptomyce
63	7	5.1	371	2	09KFO8	09KFO8 bacillus ha
64	7	5.1	374	2	09LC69	09LC69 bacillus su
65	7	5.1	379	12	070792	070792 calanthe m1
66	7	5.1	381	10	09SEV5	09SEV5 gulliardia
67	7	5.1	391	12	098Y20	098Y20 human immun
68	7	5.1	398	2	09WZL9	09WZL9 thermotoga
69	7	5.1	399	5	044999	044999 caenorhabdi
70	7	5.1	405	12	088563	088563 turnip mosa
71	7	5.1	408	12	099251	099251 turnip yell
72	7	5.1	449	5	020198	020198 caenorhabdi
73	7	5.1	476	12	09WU00	09WU00 human immun
74	7	5.1	492	12	09PYA9	09PYA9 turnip mosa
75	7	5.1	497	12	099K04	099K04 turnip mosa
76	7	5.1	499	12	09IC19	09IC19 turnip mosa
77	7	5.1	499	12	09IC18	09IC18 turnip mosa
78	7	5.1	499	12	09EAA9	09EAA9 turnip mosa
79	7	5.1	499	12	0998R4	0998R4 turnip mosa
80	7	5.1	560	12	096347	096347 human immun
81	7	5.1	563	12	09Q8G8	09Q8G8 myxoma viru
82	7	5.1	589	12	09DII9	09DII9 human immun
83	7	5.1	598	12	09DII6	09DII6 human immun
84	7	5.1	694	1	09V0H3	09V0H3 pyrococcus
85	7	5.1	721	2	086171	086171 clostridium
86	7	5.1	801	12	041344	041344 turnip mosa
87	7	5.1	999	12	090082	090082 human immun
88	7	5.1	999	12	090082	090082 human immun
89	7	5.1	1008	5	077391	077391 plasmodium
90	7	5.1	1035	12	073194	073194 human immun
91	7	5.1	1055	12	09YRU0	09YRU0 human immun
92	7	5.1	1059	12	076630	076630 human immun

93	7	5.1	1188	4	Q9H288	Q9H288 homo sapien	166	6	4.4	219	13	Q99370	Q99370 gallus gall
94	7	5.1	1599	11	Q99NH0	Q99nh0 mus musculu	167	6	4.4	226	2	Q9KBX0	Q9Kbx0 bacillus na
95	7	5.1	3164	12	Q9JIC12	Q9jic12 tunip mosa	168	6	4.4	226	11	Q9JIX3	Q9jix3 rattus norv
96	6	4.4	21	6	Q9TR36	Q9tr36 bos taurus	169	6	4.4	227	3	Q9J775	Q9j775 schizosacch
97	6	4.4	72	11	P97850	P97850 rattus norv	170	6	4.4	227	5	Q9V4C7	Q9v4c7
98	6	4.4	77	2	Q9JWU8	Q9jwu8 neisseria m	171	6	4.4	227	10	Q9FUS9	Q9fus9 arabisophila
99	6	4.4	80	2	Q9KX25	Q9kx25 xanthomonas	172	6	4.4	227	10	Q9FES8	Q9fes8 streptomyc
100	6	4.4	85	2	Q9WVG1	Q9wvg1 xanthomonas	173	6	4.4	230	2	Q9Z5A2	Q9z5a2 lycopersico
101	6	4.4	89	1	Q9BJR9	Q9bjr9 exiguobacte	174	6	4.4	230	11	Q9JKB1	Q9jkb1 streptomyc
102	6	4.4	90	1	Q9HJR9	Q9hjr9 thermoplasm	175	6	4.4	230	11	Q9EOX7	Q9eox7 mus musculu
103	6	4.4	93	3	Q12497	Q12497 saccharomyc	176	6	4.4	230	13	Q9P6C7	Q9p6c7 gallus gall
104	6	4.4	100	3	P87070	P87070 laccaria bi	177	6	4.4	231	2	Q9P6C7	Q9p6c7 gallus gall
105	6	4.4	103	5	Q1S656	Q1s656 plasmodium	178	6	4.4	231	2	Q9QX6	Q9qx6 borrelia bu
106	6	4.4	111	2	Q9S085	Q9s085 borrelia bu	179	6	4.4	231	2	Q9JPK0	Q9jpk0 neisseria m
107	6	4.4	111	2	Q06696	Q06696 borrelia bu	180	6	4.4	231	2	Q9JPU9	Q9jpu9 neisseria m
108	6	4.4	112	2	Q9FLL1	Q9fll1 thermosynec	181	6	4.4	231	11	Q9CS11	Q9cs11 mus musculu
109	6	4.4	119	2	Q9KZK3	Q9kzk3 streptomyc	182	6	4.4	232	5	Q17525	Q17525
110	6	4.4	123	10	Q9XPM3	Q9xpm3 dianthus ca	183	6	4.4	234	2	Q06424	Q06424 mycobacteri
111	6	4.4	129	8	Q9BBH0	Q9bbh0 swertia voi	184	6	4.4	237	10	Q9FHR2	Q9fhr2 arabisdopsis
112	6	4.4	135	11	Q9CQN7	Q9cqn7 mus musculu	185	6	4.4	241	2	P94514	Q9ach4 bacillus su
113	6	4.4	138	2	Q9WXS5	Q9wxs5 thermotoga	186	6	4.4	241	2	Q9ACH4	Q9ach4 uncultured
114	6	4.4	141	2	Q9JS86	Q9js86 chlamydia p	187	6	4.4	241	2	Q9ACH3	Q9ach3 uncultured
115	6	4.4	144	11	Q9RIG0	Q9rig0 rattus norv	188	6	4.4	241	2	Q9ACH2	Q9ach2 uncultured
116	6	4.4	146	5	Q45175	Q45175 caenorhabdi	189	6	4.4	241	2	Q9ACH1	Q9ach1 uncultured
117	6	4.4	158	2	Q9KCB0	Q9kcb0 bacillus na	190	6	4.4	242	2	Q9ACH0	Q9ach0 uncultured
118	6	4.4	159	5	Q9VAK5	Q9vak5 drosophila	191	6	4.4	243	2	Q53416	Q53416 bradyrhizob
119	6	4.4	162	2	Q9C140	Q9c140 lactococcus	192	6	4.4	244	2	P77222	P77222 escherichia
120	6	4.4	162	9	Q9AZD6	Q9azd6 bacterioph	193	6	4.4	244	10	Q9MLW8	Q9mlw8 arabisdopsis
121	6	4.4	164	10	Q9C910	Q9c910 arabisdopsis	194	6	4.4	247	4	Q9NXZ6	Q9nxz6 homo sapien
122	6	4.4	165	10	Q65155	Q65155 arabisdopsis	195	6	4.4	248	2	P76119	P76119 escherichia
123	6	4.4	166	2	Q66981	Q66981 aquifex ae	196	6	4.4	250	5	Q9VBV1	Q9vbv1 drosophila
124	6	4.4	166	10	Q9XEB2	Q9xeb2 arabisdopsis	197	6	4.4	250	8	Q9VG16	Q9vg16 ochromonas
125	6	4.4	169	12	Q9QFZ2	Q9qfz2 human immun	198	6	4.4	252	2	Q9RWK7	Q9rwk7 delinococcus
126	6	4.4	171	11	Q9CYX8	Q9cyx8 mus musculu	199	6	4.4	253	2	P6151	P6151 vibrio chol
127	6	4.4	173	5	Q01354	Q01354 halocynthia	200	6	4.4	254	2	Q9KOF0	Q9kfo0 vibrio chol
128	6	4.4	175	2	Q910A9	Q910a9 pseudomonas	201	6	4.4	260	13	Q9PU19	Q9pu19 xenopus lae
129	6	4.4	177	10	Q9L178	Q9l178 arabisdopsis	202	6	4.4	265	3	Q12044	Q12044 saccharomyc
130	6	4.4	179	2	Q9XCNO	Q9xcno helicobacte	203	6	4.4	269	10	Q9S972	Q9s972 arabisdopsis
131	6	4.4	180	10	Q9SSY1	Q9ssy1 cucumis sat	204	6	4.4	270	2	Q9XAK5	Q9xak5 streptomyc
132	6	4.4	183	9	Q9G0Z7	Q9g0z7 bacterioph	205	6	4.4	272	2	Q9A0N8	Q9a0n8 streptococ
133	6	4.4	186	9	Q9T169	Q9t169 bacterioph	206	6	4.4	272	5	P90762	P90762 caenorhabdi
134	6	4.4	187	2	Q9EE00	Q9ee00 human immun	207	6	4.4	272	12	Q9EHA6	Q9eha6 human immun
135	6	4.4	187	2	Q50716	Q50716 borrelia bu	208	6	4.4	273	4	Q9HD12	Q9hd12 homo sapien
136	6	4.4	192	10	Q9STF8	Q9stf8 arabisdopsis	209	6	4.4	273	12	Q9JSE1	Q9jse1 fowlpox vir
137	6	4.4	193	4	Q14426	Q14426 homo sapien	210	6	4.4	274	10	Q9T083	Q9t083 arabisdopsis
138	6	4.4	194	10	Q9SY68	Q9sy68 arabisdopsis	211	6	4.4	275	2	Q50711	Q50711 mycobacteri
139	6	4.4	195	2	Q9WXG4	Q9wxg4 rhodococcus	212	6	4.4	275	2	Q9RBF3	Q9rbf3 alcaligenes
140	6	4.4	195	2	Q9PM27	Q9pm27 campylobact	213	6	4.4	275	2	Q9Z821	Q9z821 chlamydia p
141	6	4.4	200	2	Q9CJ42	Q9cj42 lactococcus	214	6	4.4	275	11	Q9KJL6	Q9kjl6 bacillus ha
142	6	4.4	200	4	Q14426	Q14426 homo sapien	215	6	4.4	275	4	Q9N1E2	Q9n1e2 mus musculu
143	6	4.4	201	1	Q9HSE2	Q9hse2 halobacteri	216	6	4.4	276	4	Q9NT55	Q9nt55 homo sapien
144	6	4.4	204	5	Q9NSZ8	Q9nsz8 caenorhabdi	217	6	4.4	277	2	Q9ZK43	Q9zk43 helicobacte
145	6	4.4	205	12	Q88288	Q88288 snakehead r	218	6	4.4	278	5	Q01411	Q01411 teladorsagi
146	6	4.4	205	12	Q9IMT6	Q9imt6 lassa virus	219	6	4.4	278	5	Q76646	Q76646 haemonchus
147	6	4.4	209	12	Q9IMT5	Q9imt5 lassa virus	220	6	4.4	278	5	Q01410	Q01410 teladorsagi
148	6	4.4	209	12	Q9IMT5	Q9imt5 lassa virus	221	6	4.4	278	5	Q9GNP7	Q9gnp7 caenorhabdi
149	6	4.4	209	12	Q9IMT4	Q9imt4 lassa virus	222	6	4.4	279	2	Q9ZB12	Q9zb12 leucostoc
150	6	4.4	209	12	Q9IMT2	Q9imt2 lassa virus	223	6	4.4	279	2	P696923	P696923 mycobacteri
151	6	4.4	209	12	Q9IMT1	Q9imt1 lassa virus	224	6	4.4	279	2	Q9P0U8	Q9p0u8 ureaplasma
152	6	4.4	210	2	Q06405	Q06405 mycobacteri	225	6	4.4	282	4	Q9BTU2	Q9btu2 homo sapien
153	6	4.4	210	12	Q9IMT3	Q9imt3 lassa virus	226	6	4.4	279	12	Q9WBS3	Q9wbs3 physalis se
154	6	4.4	210	12	Q9IMT0	Q9imt0 lassa virus	227	6	4.4	279	12	Q9JH60	Q9jh60 melon yello
155	6	4.4	210	12	Q9IMH9	Q9imh9 lassa virus	228	6	4.4	280	5	Q25597	Q25597 onchocerca
156	6	4.4	210	12	Q9IMH2	Q9imh2 lassa virus	229	6	4.4	280	5	Q9NGY0	Q9ngy0 druglia mala
157	6	4.4	210	12	Q9IMD4	Q9imd4 lassa virus	230	6	4.4	281	5	Q9NGY1	Q9ngy1 diflorilaria
158	6	4.4	211	2	Q9EYL7	Q9eyl7 listeria mo	231	6	4.4	282	5	Q16558	Q16558 caenorhabdi
159	6	4.4	211	2	Q9AK92	Q9ak92 streptomyc	232	6	4.4	286	12	Q9JGU0	Q9jgu0 northern ce
160	6	4.4	211	11	Q55027	Q55027 mus musculu	233	6	4.4	290	2	Q87563	Q87563 bacillus fi
161	6	4.4	215	5	Q21354	Q21354 caenorhabdi	234	6	4.4	292	2	Q25452	Q25452 helicobacte
162	6	4.4	215	5	Q9UFG1	Q9ufg1 homo sapien	235	6	4.4	292	2	Q9ZL86	Q9zl86 helicobacte
163	6	4.4	216	5	Q17526	Q17526 caenorhabdi	236	6	4.4	292	4	Q9H0V8	Q9h0v8 homo sapien
164	6	4.4	218	2	Q32851	Q32851 mycobacteri	237	6	4.4	295	2	Q69877	Q69877 streptomyc
165	6	4.4	218	2	Q32851	Q32851 mycobacteri	238	6	4.4	295	5	Q9VIX0	Q9vix0 drosophila



239	6	4.4	297	1	026503	026503 methanobact	312	6	4.4	344	12	09QM05	09qm05 influenza b
240	6	4.4	299	2	09ZMX9	09zmx9 helicobacte	313	6	4.4	345	12	09QM02	09qm02 influenza b
241	6	4.4	299	2	09X440	09x440 helicobacte	314	6	4.4	345	12	082638	082638 influenza b
242	6	4.4	299	2	09X439	09x439 helicobacte	315	6	4.4	345	12	082640	082640 influenza b
243	6	4.4	299	2	09X435	09x435 helicobacte	316	6	4.4	345	12	082641	082641 influenza b
244	6	4.4	300	2	09KGN3	09kgn3 bacillus ha	317	6	4.4	345	12	082642	082642 influenza b
245	6	4.4	301	2	09KZP5	09kzp5 streptomyce	318	6	4.4	345	12	082643	082643 influenza b
246	6	4.4	301	2	052494	052494 burkholderi	319	6	4.4	345	12	082645	082645 influenza b
247	6	4.4	302	2	09K7X6	09k7x6 bacillus ha	320	6	4.4	345	12	082647	082647 influenza b
248	6	4.4	304	5	045705	045705 caenorhabdi	321	6	4.4	345	12	082648	082648 influenza b
249	6	4.4	307	5	09NMD3	09nmd3 dirosophila	322	6	4.4	345	12	082652	082652 influenza b
250	6	4.4	308	5	09VHC8	09vhc8 dirosophila	323	6	4.4	345	12	082653	082653 influenza b
251	6	4.4	315	2	09Z489	09z489 listeria iv	324	6	4.4	345	12	082659	082659 influenza b
252	6	4.4	315	12	09Q0B8	09q0b8 yaba monkey	325	6	4.4	345	12	082655	082655 influenza b
253	6	4.4	316	2	09KTJ3	09ktj3 vibrio chol	326	6	4.4	345	12	082655	082655 influenza b
254	6	4.4	316	12	09J587	09j587 fowlpox vir	327	6	4.4	345	12	09YWM7	09ywj7 influenza b
255	6	4.4	316	12	09DH01	09dh01 yaba-like d	328	6	4.4	345	12	09YWM7	09ywj7 influenza b
256	6	4.4	321	2	09KCO2	09kco2 bacillus ha	329	6	4.4	345	12	09YWM6	09ywj6 influenza b
257	6	4.4	322	2	09KCB6	09kcb6 bacillus ha	330	6	4.4	345	12	09YWM5	09ywj5 influenza b
258	6	4.4	323	2	09KCG4	09kcg4 rickettsia	331	6	4.4	345	12	09YWM4	09ywj4 influenza b
259	6	4.4	323	5	044200	044200 bombx mori	332	6	4.4	345	12	09YWM3	09ywj3 influenza b
260	6	4.4	324	5	001479	001479 caenorhabdi	333	6	4.4	345	12	09YWM2	09ywj2 influenza b
261	6	4.4	326	2	09ZEK9	09zek9 helicobacte	334	6	4.4	345	12	09YWM1	09ywj1 influenza b
262	6	4.4	328	12	096697	096697 canine herp	335	6	4.4	345	12	09YWM0	09ywj0 influenza b
263	6	4.4	328	12	09QEM3	09qem3 human immun	336	6	4.4	345	12	09YMT9	09ymj9 influenza b
264	6	4.4	328	12	09EKD3	09ekd3 human immun	337	6	4.4	345	12	09YMT8	09ymj8 influenza b
265	6	4.4	328	12	09EK53	09ek53 human immun	338	6	4.4	345	12	09YMT7	09ymj7 influenza b
266	6	4.4	328	12	09E344	09e344 human immun	339	6	4.4	345	12	09YMT6	09ymj6 influenza b
267	6	4.4	328	12	09E1Z1	09e1z1 human immun	340	6	4.4	345	12	09YMT5	09ymj5 influenza b
268	6	4.4	328	12	09ETB3	09etb3 human immun	341	6	4.4	345	12	09YMT4	09ymj4 influenza b
269	6	4.4	328	12	09ETB1	09etb1 human immun	342	6	4.4	345	12	09YMT3	09ymj3 influenza b
270	6	4.4	328	12	09ETH0	09eth0 human immun	343	6	4.4	345	12	09YMT2	09ymj2 influenza b
271	6	4.4	328	12	09EH99	09eh99 human immun	344	6	4.4	345	12	09YMT1	09ymj1 influenza b
272	6	4.4	328	12	09EH92	09eh92 human immun	345	6	4.4	345	12	09YMT0	09ymj0 influenza b
273	6	4.4	328	12	09EH88	09eh88 human immun	346	6	4.4	345	12	09YMN9	09ymj9 influenza b
274	6	4.4	328	12	09EH87	09eh87 human immun	347	6	4.4	345	12	09YMN8	09ymj8 influenza b
275	6	4.4	328	12	09EH77	09eh77 human immun	348	6	4.4	345	12	09YMN7	09ymj7 influenza b
276	6	4.4	328	12	09EH75	09eh75 human immun	349	6	4.4	345	12	09YMN6	09ymj6 influenza b
277	6	4.4	328	12	09EH69	09eh69 human immun	350	6	4.4	345	12	09YMN5	09ymj5 influenza b
278	6	4.4	328	12	09EH68	09eh68 human immun	351	6	4.4	345	12	09YMN4	09ymj4 influenza b
279	6	4.4	328	12	09EH66	09eh66 human immun	352	6	4.4	345	12	09YMN3	09ymj3 influenza b
280	6	4.4	328	12	09EH65	09eh65 human immun	353	6	4.4	345	12	09YMN2	09ymj2 influenza b
281	6	4.4	328	12	09EH63	09eh63 human immun	354	6	4.4	345	12	09YMN1	09ymj1 influenza b
282	6	4.4	328	12	09EH62	09eh62 human immun	355	6	4.4	345	12	09YMN0	09ymj0 influenza b
283	6	4.4	328	12	09EH61	09eh61 human immun	356	6	4.4	345	12	09YMG9	09ymj9 influenza b
284	6	4.4	328	12	09EH60	09eh60 human immun	357	6	4.4	345	12	09YMG8	09ymj8 influenza b
285	6	4.4	328	12	09EH57	09eh57 human immun	358	6	4.4	345	12	09YI18	09yi18 influenza b
286	6	4.4	328	12	09EH56	09eh56 human immun	359	6	4.4	345	12	09YMDB3	09ydb3 influenza b
287	6	4.4	328	12	09EHV0	09ehv0 human immun	360	6	4.4	345	12	09YMDB2	09ydb2 influenza b
288	6	4.4	328	12	09EH12	09eh12 human immun	361	6	4.4	345	12	09YMDA9	09yda9 influenza b
289	6	4.4	328	12	09ED27	09ed27 human immun	362	6	4.4	345	12	09QMU1	09qmu1 influenza b
290	6	4.4	328	12	09ED23	09ed23 human immun	363	6	4.4	345	12	09QMT7	09qmt7 influenza b
291	6	4.4	328	12	09EDW8	09edw8 human immun	364	6	4.4	345	12	09QMS4	09qms4 influenza b
292	6	4.4	328	12	09EDW5	09edw5 human immun	365	6	4.4	345	12	09YMG7	09ymj7 influenza b
293	6	4.4	328	12	09EDK1	09edk1 human immun	366	6	4.4	346	12	09YMG6	09ymj6 influenza b
294	6	4.4	328	12	09EDG0	09edg0 human immun	367	6	4.4	346	12	09YMG5	09ymj5 influenza b
295	6	4.4	328	12	09EDAO	09eda0 human immun	368	6	4.4	346	12	09YMG4	09ymj4 influenza b
296	6	4.4	328	12	09EDC3	09ecc3 human immun	369	6	4.4	346	12	09YMG3	09ymj3 influenza b
297	6	4.4	331	5	020774	020774 caenorhabdi	370	6	4.4	346	12	09YMG2	09ymj2 influenza b
298	6	4.4	331	12	09OAV7	09oav7 influenza b	371	6	4.4	346	12	09YMG1	09ymj1 influenza b
299	6	4.4	332	5	044895	044895 caenorhabdi	372	6	4.4	346	12	09YMG0	09ymj0 influenza b
300	6	4.4	332	2	053339	053339 rhizobium m	373	6	4.4	346	12	09YMF9	09ymf9 influenza b
301	6	4.4	334	10	09LW03	09lww3 arabidopsis	374	6	4.4	346	12	09YMF8	09ymf8 influenza b
302	6	4.4	335	12	09OAV6	09oav6 influenza b	375	6	4.4	346	12	09YMF7	09ymf7 influenza b
303	6	4.4	335	12	09OAV5	09oav5 influenza b	376	6	4.4	346	12	09YMF6	09ymf6 influenza b
304	6	4.4	336	12	09WPR9	09wpr9 influenza b	377	6	4.4	346	12	09YMF5	09ymf5 influenza b
305	6	4.4	337	5	09VNH7	09vnh7 dirosophila	378	6	4.4	346	12	09YMF4	09ymf4 influenza b
306	6	4.4	338	2	09AJU0	09ajno kurzthia sp.	379	6	4.4	346	12	09YMF3	09ymf3 influenza b
307	6	4.4	340	12	098YT4	098yt4 human immun	380	6	4.4	346	12	09YMF2	09ymf2 influenza b
308	6	4.4	341	12	09BWR1	09bwr1 human immun	381	6	4.4	346	12	09YMF1	09ymf1 influenza b
309	6	4.4	343	2	09RVS6	09rvs6 delinococcus	382	6	4.4	346	12	09YWF1	09ywf1 influenza b
310	6	4.4	344	12	09WDB1	09wdb1 influenza b	383	6	4.4	346	12	09YWF0	09ywf0 influenza b
311	6	4.4	344	12	09WDB0	09wdb0 influenza b	384	6	4.4	346	12	09YWE9	09ywe9 influenza b

385	6	4.4	346	12	Q9YWE8	Q9YWE8 influenza b	458	6	4.4	348	12	Q9OAV4	Q9gav4 influenza b
386	6	4.4	346	12	Q9YWE7	Q9YWE7 influenza b	459	6	4.4	348	12	Q9OAV3	Q9gav3 influenza b
387	6	4.4	346	12	Q9YWE6	Q9YWE6 influenza b	460	6	4.4	348	12	Q9OAV5	Q9gav5 influenza b
388	6	4.4	346	12	Q9YWE5	Q9YWE5 influenza b	461	6	4.4	349	12	Q9I6V5	Q9i6v5 aquifex aeo
389	6	4.4	346	12	Q9YWE4	Q9YWE4 influenza b	462	6	4.4	349	12	Q9WSR7	Q9wsr7 influenza b
390	6	4.4	346	12	Q9YWE3	Q9YWE3 influenza b	463	6	4.4	349	12	Q9I6V6	Q9i6v6 influenza b
391	6	4.4	346	12	Q9YWE2	Q9YWE2 influenza b	464	6	4.4	350	2	Q9RZ91	Q9rzt91 delinococcus
392	6	4.4	346	12	Q9YWE1	Q9YWE1 influenza b	465	6	4.4	350	2	Q9EVJ8	Q9evj8 corynebacte
393	6	4.4	346	12	Q9YWE0	Q9YWE0 influenza b	466	6	4.4	350	12	Q9WSR3	Q9wsr3 influenza b
394	6	4.4	346	12	Q9YMD9	Q9YMD9 influenza b	467	6	4.4	351	4	Q9NUT8	Q9nut8 homo sapien
395	6	4.4	346	12	Q9YMD8	Q9YMD8 influenza b	468	6	4.4	351	4	Q9BV76	Q9bv76 homo sapien
396	6	4.4	346	12	Q9YMD7	Q9YMD7 influenza b	469	6	4.4	351	12	Q9WSS1	Q9wss1 influenza b
397	6	4.4	346	12	Q9YMD6	Q9YMD6 influenza b	470	6	4.4	351	12	Q9I6V1	Q9i6v1 influenza b
398	6	4.4	346	12	Q9YMD5	Q9YMD5 influenza b	471	6	4.4	352	4	Q9BTB6	Q9btb6 homo sapien
399	6	4.4	346	12	Q9YMD4	Q9YMD4 influenza b	472	6	4.4	352	12	Q9OKD9	Q9okd9 influenza b
400	6	4.4	346	12	Q9YMD3	Q9YMD3 influenza b	473	6	4.4	352	12	Q9OKD8	Q9okd8 influenza b
401	6	4.4	346	12	Q9YMD2	Q9YMD2 influenza b	474	6	4.4	352	12	Q9OKD7	Q9okd7 influenza b
402	6	4.4	346	12	Q9YMD1	Q9YMD1 influenza b	475	6	4.4	352	12	Q9I6V2	Q9i6v2 influenza b
403	6	4.4	346	12	Q9YMD0	Q9YMD0 influenza b	476	6	4.4	353	12	Q9WSS2	Q9wss2 influenza b
404	6	4.4	346	12	Q9YMC9	Q9YMC9 influenza b	477	6	4.4	353	12	Q9WSR8	Q9wsr8 influenza b
405	6	4.4	346	12	Q9YMC8	Q9YMC8 influenza b	478	6	4.4	353	12	Q9OMH0	Q9omh0 influenza b
406	6	4.4	346	12	Q9YMC7	Q9YMC7 influenza b	479	6	4.4	354	12	Q9IP14	Q9ip14 influenza b
407	6	4.4	346	12	Q9YMC6	Q9YMC6 influenza b	480	6	4.4	355	12	Q9IP12	Q9ip12 influenza b
408	6	4.4	346	12	Q9YMC5	Q9YMC5 influenza b	481	6	4.4	355	12	Q9I6V0	Q9i6v0 influenza b
409	6	4.4	346	12	Q9YMC4	Q9YMC4 influenza b	482	6	4.4	355	12	Q9WSR6	Q9wsr6 influenza b
410	6	4.4	346	12	Q9YMC3	Q9YMC3 influenza b	483	6	4.4	356	12	Q9WSR4	Q9wsr4 influenza b
411	6	4.4	346	12	Q9YMC2	Q9YMC2 influenza b	484	6	4.4	356	12	Q9I6V3	Q9i6v3 influenza b
412	6	4.4	346	12	Q9YMC1	Q9YMC1 influenza b	485	6	4.4	357	2	Q9YR85	Q9yr85 streptococc
413	6	4.4	346	12	Q9YJ12	Q9YJ12 influenza b	486	6	4.4	358	12	Q9WSR9	Q9wsr9 influenza b
414	6	4.4	346	12	Q9YJ10	Q9YJ10 influenza b	487	6	4.4	358	12	Q9WRK3	Q9wrk3 influenza b
415	6	4.4	346	12	Q9YJ09	Q9YJ09 influenza b	488	6	4.4	359	12	Q9WSS0	Q9wss0 influenza b
416	6	4.4	346	12	Q9YJ06	Q9YJ06 influenza b	489	6	4.4	359	12	Q9WSR5	Q9wsr5 influenza b
417	6	4.4	346	12	Q9YJ03	Q9YJ03 influenza b	490	6	4.4	360	12	Q9I6V3	Q9i6v3 influenza b
418	6	4.4	346	12	Q9YIY4	Q9YIY4 influenza b	491	6	4.4	360	12	Q9I6V2	Q9i6v2 influenza b
419	6	4.4	346	12	Q9YIY2	Q9YIY2 influenza b	492	6	4.4	360	12	Q9I6V0	Q9i6v0 influenza b
420	6	4.4	346	12	Q9YIY6	Q9YIY6 influenza b	493	6	4.4	360	12	Q9I6V1	Q9i6v1 influenza b
421	6	4.4	346	12	Q9YIY7	Q9YIY7 influenza b	494	6	4.4	360	12	Q9I6V3	Q9i6v3 influenza b
422	6	4.4	346	12	Q9YMD8	Q9YMD8 influenza b	495	6	4.4	360	12	Q9I6V3	Q9i6v3 influenza b
423	6	4.4	346	12	Q9YMD5	Q9YMD5 influenza b	496	6	4.4	360	12	Q9I6V3	Q9i6v3 influenza b
424	6	4.4	346	12	Q9YMD5	Q9YMD5 influenza b	497	6	4.4	360	12	Q9I6V3	Q9i6v3 influenza b
425	6	4.4	346	12	Q9YMD5	Q9YMD5 influenza b	498	6	4.4	360	12	Q9I6V3	Q9i6v3 influenza b
426	6	4.4	346	12	Q9YMD4	Q9YMD4 influenza b	499	6	4.4	360	12	Q9I6V3	Q9i6v3 influenza b
427	6	4.4	346	12	Q9YMD4	Q9YMD4 influenza b	500	6	4.4	360	12	Q9I6V3	Q9i6v3 influenza b
428	6	4.4	346	12	Q9YMD4	Q9YMD4 influenza b	501	6	4.4	360	12	Q9YX39	Q9yx39 influenza b
429	6	4.4	346	12	Q9YMD6	Q9YMD6 influenza b	502	6	4.4	360	12	Q9QMT9	Q9qmt9 influenza b
430	6	4.4	346	12	Q9YMD6	Q9YMD6 influenza b	503	6	4.4	360	12	Q9QMT8	Q9qmt8 influenza b
431	6	4.4	346	12	Q9YMD6	Q9YMD6 influenza b	504	6	4.4	360	12	Q9QMT6	Q9qmt6 influenza b
432	6	4.4	346	12	Q9YMD6	Q9YMD6 influenza b	505	6	4.4	360	12	Q9QMT5	Q9qmt5 influenza b
433	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	506	6	4.4	360	12	Q9QMT0	Q9qmt0 influenza b
434	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	507	6	4.4	361	1	Q9Y527	Q9y527 pyrococcus
435	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	508	6	4.4	361	3	Q9Y725	Q9y725 candida gla
436	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	509	6	4.4	361	12	Q9YX46	Q9yx46 influenza b
437	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	510	6	4.4	361	12	Q9YX45	Q9yx45 influenza b
438	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	511	6	4.4	361	12	Q9YX41	Q9yx41 influenza b
439	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	512	6	4.4	361	12	Q9YX43	Q9yx43 influenza b
440	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	513	6	4.4	361	12	Q9YX44	Q9yx44 influenza b
441	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	514	6	4.4	361	12	Q9YX40	Q9yx40 influenza b
442	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	515	6	4.4	361	12	Q9YX40	Q9yx40 influenza b
443	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	516	6	4.4	361	12	Q9YX40	Q9yx40 influenza b
444	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	517	6	4.4	361	12	Q9YX40	Q9yx40 influenza b
445	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	518	6	4.4	361	12	Q9YX40	Q9yx40 influenza b
446	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	519	6	4.4	361	12	Q9YX40	Q9yx40 influenza b
447	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	520	6	4.4	361	12	Q9YX40	Q9yx40 influenza b
448	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	521	6	4.4	361	12	Q9YX40	Q9yx40 influenza b
449	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	522	6	4.4	361	12	Q9YX40	Q9yx40 influenza b
450	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	523	6	4.4	361	12	Q9YX40	Q9yx40 influenza b
451	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	524	6	4.4	361	12	Q9YX40	Q9yx40 influenza b
452	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	525	6	4.4	361	12	Q9YX40	Q9yx40 influenza b
453	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	526	6	4.4	361	12	Q9YX40	Q9yx40 influenza b
454	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	527	6	4.4	361	12	Q9YX40	Q9yx40 influenza b
455	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	528	6	4.4	361	12	Q9YX40	Q9yx40 influenza b
456	6	4.4	347	12	Q9YMD6	Q9YMD6 influenza b	529	6	4.4	361	12	Q9YX40	Q9yx40 influenza b
457	6	4.4	348	2	Q9I4X1	Q9i4x1 pseudomonas	530	6	4.4	362	12	Q9I6V3	Q9i6v3 influenza b

531	6	4.4	363	12	091U5	Q91Y5 influenza b	604	6	4.4	421	1	091800	Q91800 aeropyrum p
532	6	4.4	363	12	091Y3	Q91Y3 influenza b	605	6	4.4	424	2	091232	Q91232 pseudomonas
533	6	4.4	363	12	091Y2	Q91Y2 influenza b	606	6	4.4	424	5	016594	Q16594 caenorhabdi
534	6	4.4	363	12	091Y1	Q91Y1 influenza b	607	6	4.4	424	10	095VK8	Q95VK8 arbidopsi
535	6	4.4	364	3	Q9P8N0	Q9P8N0 picchia angu	608	6	4.4	425	1	059010	059010 pyrococcus
536	6	4.4	364	12	084081	084081 influenza b	609	6	4.4	425	1	09V0E7	Q9V0E7 pyrococcus
537	6	4.4	365	4	Q9H909	Q9H909 homo sapien	610	6	4.4	426	10	Q9H06	Q9H06 halobacteri
538	6	4.4	365	10	Q9F035	Q9F035 arbidopsi	611	6	4.4	426	10	Q9S0L7	Q9S0L7 arbidopsi
539	6	4.4	366	12	Q9AVR1	Q9AVR1 stephanotis	612	6	4.4	429	5	044842	044842 caenorhabdi
540	6	4.4	366	12	084076	084076 influenza b	613	6	4.4	431	2	Q9KLR1	Q9KLR1 vibrio chol
541	6	4.4	367	1	Q9HRA3	Q9HRA3 halobacteri	614	6	4.4	433	1	028602	028602 archaeoglob
542	6	4.4	368	12	Q9WPS3	Q9WPS3 influenza b	615	6	4.4	433	5	Q9V438	Q9V438 drosophila
543	6	4.4	369	10	Q9LKH6	Q9LKH6 arbidopsi	616	6	4.4	433	5	Q9NK40	Q9NK40 drosophila
544	6	4.4	371	5	Q9N470	Q9N470 caenorhabdi	617	6	4.4	434	5	Q9VAT9	Q9VAT9 drosophila
545	6	4.4	371	12	Q98220	Q98220 human immun	618	6	4.4	435	2	Q9K832	Q9K832 bacillus ha
546	6	4.4	373	2	051151	051151 neisseria m	619	6	4.4	435	10	Q9PPA2	Q9PPA2 oryza sativ
547	6	4.4	373	6	097932	097932 ovis aries	620	6	4.4	436	2	051415	051415 boerelia bu
548	6	4.4	373	10	Q9ZYC1	Q9ZYC1 arbidopsi	621	6	4.4	436	5	Q9XTU8	Q9XTU8 caenorhabdi
549	6	4.4	374	10	Q9FRC8	Q9FRC8 arbidopsi	622	6	4.4	438	2	Q9C8R3	Q9C8R3 mycobacteri
550	6	4.4	376	2	052711	052711 saccharopol	623	6	4.4	438	5	Q20252	Q20252 caenorhabdi
551	6	4.4	376	6	046499	046499 bos taurus	624	6	4.4	438	5	Q9VJQ3	Q9VJQ3 drosophila
552	6	4.4	376	6	Q9TXX4	Q9TXX4 capra hircu	625	6	4.4	440	2	Q919C1	Q919C1 azoarcus ev
553	6	4.4	376	10	049315	049315 arbidopsi	626	6	4.4	440	5	077369	077369 plasmodium
554	6	4.4	377	2	Q9XMS5	Q9XMS5 arbidopsi	627	6	4.4	441	2	Q9R1T7	Q9R1T7 streptomyc
555	6	4.4	377	2	Q9XSE8	Q9XSE8 streptomyc	628	6	4.4	443	2	Q91IB6	Q91IB6 pseudomonas
556	6	4.4	377	3	074794	074794 schizosacch	629	6	4.4	447	2	Q9RUD3	Q9RUD3 deinococcus
557	6	4.4	377	5	Q9U053	Q9U053 giardia lam	630	6	4.4	450	2	Q9KEX2	Q9KEX2 bacillus ha
558	6	4.4	377	5	Q9XVP3	Q9XVP3 caenorhabdi	631	6	4.4	452	11	Q9CXM9	Q9CXM9 mus muscul
559	6	4.4	377	12	Q9WPS2	Q9WPS2 influenza b	632	6	4.4	453	2	Q9KNH7	Q9KNH7 mus muscul
560	6	4.4	378	12	084102	084102 influenza b	633	6	4.4	455	2	053182	053182 rhodobacter
561	6	4.4	378	12	067362	067362 influenza b	634	6	4.4	455	2	Q9S228	Q9S228 rhodobacter
562	6	4.4	378	12	067364	067364 influenza b	635	6	4.4	458	12	Q9WPS1	Q9WPS1 influenza b
563	6	4.4	378	12	067365	067365 influenza b	636	6	4.4	459	11	Q99LM3	Q99LM3 mus muscul
564	6	4.4	378	12	067367	067367 influenza b	637	6	4.4	461	11	Q9DCV1	Q9DCV1 mus muscul
565	6	4.4	378	12	067368	067368 influenza b	638	6	4.4	463	10	Q39871	Q39871 glycine max
566	6	4.4	379	5	Q9U1C2	Q9U1C2 leishmania	639	6	4.4	464	5	015827	015827 leishmania
567	6	4.4	379	5	Q9BLM4	Q9BLM4 leishmania	640	6	4.4	465	2	082877	082877 streptococ
568	6	4.4	379	12	067363	067363 influenza b	641	6	4.4	465	10	Q43166	Q43166 solanum tub
569	6	4.4	379	12	067366	067366 influenza b	642	6	4.4	466	5	Q93712	Q93712 caenorhabdi
570	6	4.4	380	12	Q9LTX8	Q9LTX8 influenza b	643	6	4.4	466	10	Q43526	Q43526 lycopersico
571	6	4.4	383	2	Q9RS15	Q9RS15 deinococcus	644	6	4.4	467	1	Q93661	Q93661 methanosarc
572	6	4.4	389	9	Q9RUJ7	Q9RUJ7 streptomyc	645	6	4.4	467	10	081143	081143 lycopersico
573	6	4.4	391	5	Q96774	Q96774 tetrahymena	646	6	4.4	468	4	Q9Y4B8	Q9Y4B8 homo sapien
574	6	4.4	392	4	Q9BU06	Q9BU06 homo sapien	647	6	4.4	468	10	081722	081722 arbidopsi
575	6	4.4	394	10	Q9FJZ1	Q9FJZ1 arbidopsi	648	6	4.4	470	11	Q62965	Q62965 rattus norv
576	6	4.4	394	12	Q9WPS4	Q9WPS4 influenza b	649	6	4.4	471	5	Q9VES7	Q9VES7 drosophila
577	6	4.4	394	12	Q9GTF9	Q9GTF9 svts2 plect	650	6	4.4	471	5	Q9VES6	Q9VES6 drosophila
578	6	4.4	396	1	Q9HS71	Q9HS71 halobacteri	651	6	4.4	474	5	Q9BRJ3	Q9BRJ3 giardia lam
579	6	4.4	397	6	Q28546	Q28546 ovis aries	652	6	4.4	476	2	Q9KLD4	Q9KLD4 vibrio chol
580	6	4.4	397	11	Q9EBJ1	Q9EBJ1 rattus norv	653	6	4.4	476	10	Q93365	Q93365 nicotiana t
581	6	4.4	397	11	Q9CWK7	Q9CWK7 mus muscul	654	6	4.4	476	10	Q9AR54	Q9AR54 nicotiana t
582	6	4.4	398	4	Q9UE68	Q9UE68 homo sapien	655	6	4.4	480	4	Q9H1P3	Q9H1P3 homo sapien
583	6	4.4	398	5	Q18884	Q18884 caenorhabdi	656	6	4.4	480	5	Q9BHJ7	Q9BHJ7 trypanosoma
584	6	4.4	399	4	Q9H7J3	Q9H7J3 homo sapien	657	6	4.4	481	5	Q22539	Q22539 caenorhabdi
585	6	4.4	402	2	Q99VE0	Q99VE0 staphylococ	658	6	4.4	482	10	082443	082443 zea mays (m
586	6	4.4	402	10	Q9LJZ8	Q9LJZ8 arbidopsi	659	6	4.4	483	2	Q56622	Q56622 vibrio chol
587	6	4.4	406	10	Q9FE30	Q9FE30 arbidopsi	660	6	4.4	483	10	087164	087164 oryza sativ
588	6	4.4	407	2	067720	067720 aquifex aeo	661	6	4.4	483	10	09S768	Q9S768 oryza sativ
589	6	4.4	407	10	Q9ZOA3	Q9ZOA3 arbidopsi	662	6	4.4	486	2	Q9L378	Q9L378 bacillus sp
590	6	4.4	408	4	Q16670	Q16670 homo sapien	663	6	4.4	486	2	Q9KW38	Q9KW38 wolbachia s
591	6	4.4	409	1	Q9YAC7	Q9YAC7 aeropyrum p	664	6	4.4	486	10	Q9M1W0	Q9M1W0 arbidopsi
592	6	4.4	410	2	Q9L446	Q9L446 streptococ	665	6	4.4	489	2	P71048	P71048 bacillus su
593	6	4.4	410	2	Q9L445	Q9L445 streptococ	666	6	4.4	493	5	Q23330	Q23330 caenorhabdi
594	6	4.4	410	2	Q9L444	Q9L444 streptococ	667	6	4.4	495	10	Q23330	Q23330 arbidopsi
595	6	4.4	410	2	Q9L388	Q9L388 streptococ	668	6	4.4	496	5	018683	018683 drosophila
596	6	4.4	410	2	Q9E258	Q9E258 corynebacte	669	6	4.4	503	2	Q9K445	Q9K445 wolbachia s
597	6	4.4	410	2	Q9ANW7	Q9ANW7 streptococ	670	6	4.4	506	10	Q38860	Q38860 arbidopsi
598	6	4.4	412	2	Q25885	Q25885 helicobacte	671	6	4.4	507	11	Q9C739	Q9C739 mus muscul
599	6	4.4	412	5	Q44190	Q44190 caenorhabdi	672	6	4.4	508	5	Q45345	Q45345 caenorhabdi
600	6	4.4	413	1	Q26879	Q26879 methanobact	673	6	4.4	508	5	Q96146	Q96146 plasmodium
601	6	4.4	414	2	053379	053379 mycobacteri	674	6	4.4	509	5	Q9VTZ7	Q9VTZ7 drosophila
602	6	4.4	416	2	Q45154	Q45154 bacteroides	675	6	4.4	509	10	P93058	P93058 brassica ju
603	6	4.4	418	10	Q9C705	Q9C705 arbidopsi	676	6	4.4	511	10	Q91F12	Q91F12 arbidopsi

677	6	4.4	513	10	Q9SD51	Q9sd51 arabidopsis	750	6	4.4	652	10	Q9ZT78	Q9ZT78 arabidopsis
678	6	4.4	513	10	P98188	P98188 vicia sativ	751	6	4.4	662	5	Q9VRV6	Q9VRV6 diosiphila
679	6	4.4	515	2	Q9RLV2	Q9rlv2 lactococcus	752	6	4.4	665	5	Q24966	Q24966 giardia lam
680	6	4.4	518	10	Q9SD15	Q9sd15 arabidopsis	753	6	4.4	672	5	Q44189	Q44189 caenorhabdi
681	6	4.4	518	11	Q9RQJ2	Q9eqj2 rattus norv	754	6	4.4	676	2	Q9PEP2	Q9PEP2 campylobact
682	6	4.4	522	10	Q9FXB5	Q9fxb5 arabidopsis	755	6	4.4	677	2	Q34323	Q34323 bacillus su
683	6	4.4	523	2	Q91729	Q91729 pseudomonas	756	6	4.4	677	4	Q9HAR0	Q9HAR0 homo sapien
684	6	4.4	526	10	Q9LZ26	Q9lzz26 arabidopsis	757	6	4.4	678	2	Q54679	Q54679 streptomyce
685	6	4.4	528	3	Q9Y7C8	Q9Y7C8 aspergillus	758	6	4.4	684	10	Q9XH27	Q9XH27 arabidopsis
686	6	4.4	529	3	Q9HDX7	Q9hdx7 schizosacch	759	6	4.4	686	5	Q96245	Q96245 plasmodium
687	6	4.4	531	5	Q9W2A9	Q9w2a9 diosiphila	760	6	4.4	691	5	Q9NKR7	Q9NKR7 leishmania
688	6	4.4	531	10	Q42582	Q42582 arabidopsis	761	6	4.4	692	2	Q9F2J1	Q9F2J1 streptomyce
689	6	4.4	531	10	Q9S1Z3	Q9s1z3 arabidopsis	762	6	4.4	694	2	Q9L5M3	Q9L5M3 salmonella
690	6	4.4	532	2	Q9LCB8	Q9lcb8 streptococc	763	6	4.4	700	4	Q9UPF8	Q9UPF8 homo sapien
691	6	4.4	533	2	Q9T069	Q9t069 arabidopsis	764	6	4.4	704	5	Q02059	Q02059 caenorhabdi
692	6	4.4	533	2	Q9U092	Q9u092 staphylococ	765	6	4.4	706	4	Q9H7N2	Q9H7N2 homo sapien
693	6	4.4	534	5	Q9U211	Q9u211 caenorhabdi	766	6	4.4	706	11	Q9D2M5	Q9D2M5 mus musculu
694	6	4.4	535	10	Q49600	Q49600 caenorhabdi	767	6	4.4	716	10	Q9FH85	Q9FH85 arabidopsis
695	6	4.4	535	10	Q9ASV4	Q9asv4 arabidopsis	768	6	4.4	717	4	Q9NX10	Q9NX10 homo sapien
696	6	4.4	536	4	Q9NQ09	Q9nq09 homo sapien	769	6	4.4	720	2	Q9X8G8	Q9X8G8 streptomyce
697	6	4.4	537	10	Q9FZ24	Q9fz24 arabidopsis	770	6	4.4	732	5	Q9VY71	Q9VY71 diosiphila
698	6	4.4	538	2	Q9PNT5	Q9pnt5 campylobact	771	6	4.4	732	10	Q04269	Q04269 chlamydomon
699	6	4.4	540	2	Q31716	Q31716 bacillus su	772	6	4.4	737	5	Q17305	Q17305 caenorhabdi
700	6	4.4	542	2	Q50025	Q50025 mycobacteri	773	6	4.4	738	2	Q59490	Q59490 lactobacill
701	6	4.4	542	5	Q9U1H6	Q9u1h6 diosiphila	774	6	4.4	738	13	Q985V5	Q985V5 xenopus lae
702	6	4.4	543	5	Q76521	Q76521 plasmodium	775	6	4.4	750	6	Q9GK13	Q9GK13 bos taurus
703	6	4.4	545	5	Q44019	Q44019 plasmodium	776	6	4.4	753	5	Q44861	Q44861 caenorhabdi
704	6	4.4	545	11	Q08661	Q08661 rattus norv	777	6	4.4	756	10	Q22863	Q22863 arabidopsis
705	6	4.4	546	5	Q9GUC3	Q9guc3 caenorhabdi	778	6	4.4	757	1	Q9H5N5	Q9H5N5 halobacteri
706	6	4.4	548	2	Q9RMQ9	Q9rmq9 delnoccocus	779	6	4.4	760	2	Q9RKG0	Q9RKG0 ashbya goss
707	6	4.4	550	2	Q9CIL3	Q9cil3 lactococcus	780	6	4.4	761	3	Q9HF59	Q9HF59 arabidopsis
708	6	4.4	554	12	Q03259	Q03259 chimpanzee	781	6	4.4	764	10	Q9Z067	Q9Z067 plasmodium
709	6	4.4	557	10	Q9FLK3	Q9flk3 arabidopsis	782	6	4.4	765	5	Q26018	Q26018 plasmodium
710	6	4.4	560	12	Q99AV8	Q99av8 human immun	783	6	4.4	767	4	Q75166	Q75166 homo sapien
711	6	4.4	562	12	Q9DW57	Q9dw57 rat cytoleg	784	6	4.4	767	5	Q18163	Q18163 caenorhabdi
712	6	4.4	563	11	Q9EQJ3	Q9eqj3 rattus norv	785	6	4.4	768	5	Q9RTE3	Q9RTE3 delnoccocus
713	6	4.4	565	2	Q9KML1	Q9kml1 vibrio chol	786	6	4.4	771	5	Q22783	Q22783 caenorhabdi
714	6	4.4	565	6	Q97582	Q97582 sus scrofa	787	6	4.4	772	2	Q9RYQ0	Q9RYQ0 delnoccocus
715	6	4.4	565	10	Q9LRT9	Q9lrt9 arabidopsis	788	6	4.4	772	5	Q60958	Q60958 leishmania
716	6	4.4	567	12	Q9LW63	Q9lw63 influenza b	789	6	4.4	781	10	Q9SF87	Q9SF87 arabidopsis
717	6	4.4	568	10	Q9LOA2	Q9loa2 arabidopsis	790	6	4.4	782	5	Q20373	Q20373 caenorhabdi
718	6	4.4	569	2	Q20817	Q20817 acetobacter	791	6	4.4	794	2	Q9X212	Q9X212 thernotoga
719	6	4.4	569	12	Q9IMT7	Q9imt7 influenza b	792	6	4.4	796	10	Q9FT77	Q9FT77 arabidopsis
720	6	4.4	569	12	Q9IMT7	Q9imt7 influenza b	793	6	4.4	800	3	Q59960	Q59960 saccharomyc
721	6	4.4	570	12	Q98102	Q98102 influenza b	794	6	4.4	805	5	Q9N5N5	Q9N5N5 euplates ae
722	6	4.4	570	12	Q98103	Q98103 influenza b	795	6	4.4	806	10	Q9FFZ5	Q9FFZ5 arabidopsis
723	6	4.4	579	11	Q9JMK5	Q9jmk5 rattus norv	796	6	4.4	807	5	Q9GZ12	Q9GZ12 caenorhabdi
724	6	4.4	579	11	Q9EQJ4	Q9eqj4 rattus norv	797	6	4.4	821	3	Q308581	Q308581 caenorhabdi
725	6	4.4	580	2	Q9KML4	Q9kml4 vibrio chol	798	6	4.4	822	1	Q30286	Q30286 archaeyoc
726	6	4.4	581	3	Q14035	Q14035 schizosacch	799	6	4.4	832	5	Q25751	Q25751 mus musculu
727	6	4.4	582	12	Q84097	Q84097 influenza b	800	6	4.4	837	10	Q9LZC7	Q9LZC7 plasmodium
728	6	4.4	584	12	Q9W9R2	Q9w9r2 influenza b	801	6	4.4	837	10	Q9LZC7	Q9LZC7 arabidopsis
729	6	4.4	588	10	Q9EFC0	Q9efc0 arabidopsis	802	6	4.4	839	10	Q9W332	Q9W332 arabidopsis
730	6	4.4	589	2	Q9REC7	Q9rec7 clostridium	803	6	4.4	844	5	Q44476	Q44476 caenorhabdi
731	6	4.4	589	13	Q98UK1	Q98uk1 carassius a	804	6	4.4	856	10	Q9MB88	Q9MB88 brassica ol
732	6	4.4	591	2	Q9FIR9	Q9fir9 clostridium	805	6	4.4	860	3	Q94543	Q94543 archais hyp
733	6	4.4	591	5	Q9GU24	Q9gu24 entamoeba h	806	6	4.4	862	10	Q94543	Q94543 schizosacch
734	6	4.4	597	10	Q9LUV9	Q9luy9 arabidopsis	807	6	4.4	865	2	Q43819	Q43819 aeromonas c
735	6	4.4	598	2	Q51954	Q51954 arabidopsis	808	6	4.4	865	5	Q9F9Q8	Q9F9Q8 aeromonas h
736	6	4.4	598	2	Q9XPD2	Q9xpd2 enterobacte	809	6	4.4	865	5	Q9V5Z4	Q9V5Z4 diosiphila
737	6	4.4	600	11	Q55185	Q55185 mus musculu	810	6	4.4	865	13	Q98SN6	Q98SN6 phaseolus v
738	6	4.4	603	4	Q9HBD7	Q9hbd7 homo sapien	811	6	4.4	865	10	Q9ZUT6	Q9ZUT6 gallus gall
739	6	4.4	604	10	Q9LW37	Q9lw37 arabidopsis	812	6	4.4	870	10	Q9LW37	Q9LW37 arabidopsis
740	6	4.4	604	10	Q9FIR8	Q9fir8 arabidopsis	813	6	4.4	874	2	Q53493	Q53493 mycobacteri
741	6	4.4	610	5	Q96198	Q96198 arabidopsis	814	6	4.4	876	10	Q42846	Q42846 hordeum vul
742	6	4.4	610	10	Q9ERT8	Q9ert8 plasmodium	815	6	4.4	883	10	Q9LRP0	Q9LRP0 caenorhabdi
743	6	4.4	612	2	Q99YI8	Q99yi8 streptococc	816	6	4.4	887	5	Q19469	Q19469 caenorhabdi
744	6	4.4	614	9	Q37967	Q37967 bacterioph	817	6	4.4	892	2	Q55676	Q55676 synechocyst
745	6	4.4	631	10	Q45820	Q45820 clostridium	818	6	4.4	896	10	Q9FEN1	Q9FEN1 arabidopsis
746	6	4.4	635	2	Q45820	Q45820 pseudomonas	819	6	4.4	898	11	Q88193	Q88193 mus musculu
747	6	4.4	640	2	Q9C5U9	Q9c5u9 pseudomonas	820	6	4.4	901	11	Q88193	Q88193 mus musculu
748	6	4.4	644	2	Q9CGM3	Q9cg93 lactococcus	821	6	4.4	902	5	Q9VSAS	Q9VSAS diosiphila
749	6	4.4	652	5	Q9U0D4	Q9u0d4 plasmodium	822	6	4.4				

823	6	4.4	902	10	082323	082323 arabidopsis	896	6	4.4	1629	10	004698	004698 pisum sativ
824	6	4.4	909	11	088190	088190 mus musculus	897	6	4.4	1638	4	09y6j4	09y6j4 homo sapien
825	6	4.4	910	9	037959	037959 lactococcus	898	6	4.4	1650	5	077328	077328 plasmodium
826	6	4.4	913	11	088195	088195 mus musculus	899	6	4.4	1712	4	09y6j5	09y6j5 homo sapien
827	6	4.4	914	5	09BMN6	09BMN6 mus musculus	900	6	4.4	1721	5	09UAR3	09UAR3 pacifastacu
828	6	4.4	914	11	088192	088192 mus musculus	901	6	4.4	1742	2	055583	055583 synechocyst
829	6	4.4	914	11	088194	088194 mus musculus	902	6	4.4	1754	4	09U1W2	09U1W2 homo sapien
830	6	4.4	916	4	060502	060502 homo sapien	903	6	4.4	1780	5	09VUC2	09VUC2 drosophila
831	6	4.4	925	5	09N8G2	09N8G2 trypanosoma	904	6	4.4	1793	10	09SJP0	09SJP0 arabidopsis
832	6	4.4	936	2	09KJ17	09KJ17 trichosemi	905	6	4.4	1787	8	09BBN6	09BBN6 lotus japon
833	6	4.4	941	4	09P2S6	09P2S6 homo sapien	906	6	4.4	1872	11	P70208	P70208 mus musculus
834	6	4.4	946	11	088690	088690 mus musculus	907	6	4.4	1883	2	09PFR2	09PFR2 ureaplasma
835	6	4.4	947	5	008667	008667 plasmodium	908	6	4.4	1888	11	088466	088466 mus musculus
836	6	4.4	947	11	088689	088689 mus musculus	909	6	4.4	1894	11	P70206	P70206 mus musculus
837	6	4.4	949	4	09Y4G6	09Y4G6 homo sapien	910	6	4.4	1905	13	091823	091823 xenopus lae
838	6	4.4	961	10	065230	065230 arabidopsis	911	6	4.4	1948	11	028189	028189 archaeoglob
839	6	4.4	969	5	017795	017795 caenorhabdi	912	6	4.4	1956	11	062968	062968 rattus norv
840	6	4.4	976	2	09EV17	09EV17 actinomyces	913	6	4.4	1957	11	063554	063554 rattus norv
841	6	4.4	984	2	067280	067280 aquilex aeo	914	6	4.4	1958	11	P70276	P70276 mus musculus
842	6	4.4	987	3	09USU1	09USU1 schizosacch	915	6	4.4	1968	5	020439	020439 caenorhabdi
843	6	4.4	987	10	09LNC5	09LNC5 arabidopsis	916	6	4.4	2070	10	09MAV3	09MAV3 arabidopsis
844	6	4.4	988	5	024418	024418 drosophila	917	6	4.4	2119	5	09VAV5	09VAV5 drosophila
845	6	4.4	997	5	024418	024418 drosophila	918	6	4.4	2172	10	048534	048534 arabidopsis
846	6	4.4	1000	6	028737	028737 oryctolagus	919	6	4.4	2206	5	096205	096205 plasmodium
847	6	4.4	1025	5	025693	025693 plasmodium	920	6	4.4	2209	5	097324	097324 plasmodium
848	6	4.4	1025	13	09DEU1	09DEU1 brachydanio	921	6	4.4	2232	5	P91365	P91365 caenorhabdi
849	6	4.4	1029	4	09C099	09C099 homo sapien	922	6	4.4	2231	5	022847	022847 caenorhabdi
850	6	4.4	1030	10	09C8T4	09C8T4 arabidopsis	923	6	4.4	2340	2	09ZD91	09ZD91 rickettsia
851	6	4.4	1039	10	09FNG8	09FNG8 arabidopsis	924	6	4.4	2485	5	096134	096134 plasmodium
852	6	4.4	1031	2	09MXJ2	09MXJ2 psychromona	925	6	4.4	2567	5	09U076	09U076 plasmodium
853	6	4.4	1036	1	029316	029316 archaeoglob	926	6	4.4	2841	2	09FBB3	09FBB3 streptomyc
854	6	4.4	1047	2	09RB35	09RB35 cytophaga s	927	6	4.4	3112	5	09NKP1	09NKP1 leishmania
855	6	4.4	1057	2	09CKR5	09CKR5 pasteurilla	928	6	4.4	3198	5	09U8G8	09U8G8 manduca sex
856	6	4.4	1081	5	09U631	09U631 drosophila	929	6	4.4	3394	4	09Y6V0	09Y6V0 homo sapien
857	6	4.4	1083	5	09VTT0	09VTT0 drosophila	930	6	4.4	3488	5	P91257	P91257 caenorhabdi
858	6	4.4	1088	10	09F1M1	09F1M1 arabidopsis	931	6	4.4	3978	5	097236	097236 plasmodium
859	6	4.4	1096	10	09LXD9	09LXD9 arabidopsis	932	6	4.4	4150	2	09KIT4	09KIT4 streptomyc
860	6	4.4	1103	3	09C2J7	09C2J7 neurospora	933	6	4.4	4345	2	09X4W2	09X4W2 vibrio chol
861	6	4.4	1107	5	09GR76	09GR76 leishmania	934	6	4.4	4547	5	09W343	09W343 drosophila
862	6	4.4	1127	12	0293K9	0293K9 caallitrichi	935	6	4.4	4558	2	09KS12	09KS12 vibrio chol
863	6	4.4	1140	5	020299	020299 caenorhabdi	936	6	4.4	4589	5	076506	076506 tetrahymena
864	6	4.4	1189	10	09FNU0	09FNU0 arabidopsis	937	6	4.4	4688	2	09P008	09P008 drosophila
865	6	4.4	1235	1	026597	026597 methanobact	938	6	4.4	4899	5	09VR81	09VR81 ureaplasma
866	6	4.4	1244	4	000562	000562 homo sapien	939	6	4.4	4967	4	092736	092736 homo sapien
867	6	4.4	1262	4	09B873	09B273 homo sapien	940	6	4.4	4967	11	09ERN6	09ERN6 mus musculus
868	6	4.4	1264	4	020684	020684 caenorhabdi	941	6	4.4	4968	6	029621	029621 oryctolagus
869	6	4.4	1270	5	062462	062462 caenorhabdi	942	6	4.4	4981	5	077372	077372 plasmodium
870	6	4.4	1272	10	09M197	09M197 arabidopsis	943	6	4.4	5458	5	09U459	09U459 plasmodium
871	6	4.4	1274	10	09YOK2	09YOK2 arabidopsis	944	6	4.4	6875	6	028733	028733 oryctolagus
872	6	4.4	1276	10	081059	081059 arabidopsis	945	6	4.4	26926	4	010466	010466 homo sapien
873	6	4.4	1284	2	09POA1	09POA1 ureaplasma	946	6	4.4	7	2	047029	047029 enterobacte
874	6	4.4	1291	10	09SFE1	09SFE1 arabidopsis	947	6	3.7	20	10	09S8J5	09S8J5 psophocarpu
875	6	4.4	1300	4	013999	013999 homo sapien	948	6	3.7	22	12	084254	084254 bovine papl
876	6	4.4	1330	10	09C536	09C536 arabidopsis	949	6	3.7	24	10	P82139	P82139 splnacia ol
877	6	4.4	1337	11	061595	061595 mus musculu	950	6	3.7	26	4	09P2U0	09P2U0 homo sapien
878	6	4.4	1330	6	097961	097961 vulpes vulp	951	6	3.7	26	5	09NBB1	09NBB1 drosophila
879	6	4.4	1352	10	09SXB2	09SXB2 arabidopsis	952	6	3.7	26	5	09BLZ7	09BLZ7 mactirotache
880	6	4.4	1352	10	09M2D1	09M2D1 arabidopsis	953	6	3.7	26	5	09BLZ6	09BLZ6 adineta vag
881	6	4.4	1352	10	09C739	09C739 arabidopsis	954	6	3.7	26	8	P92646	P92646 lialis jica
882	6	4.4	1354	4	013464	013464 homo sapien	955	6	3.7	26	8	P92772	P92772 xantusia vi
883	6	4.4	1354	6	077819	077819 oryctolagus	956	6	3.7	26	11	09QXG9	09QXG9 rattus sp.
884	6	4.4	1354	11	P70335	P70335 mus musculu	957	6	3.7	28	11	09QXB4	09QXB4 mus musculu
885	6	4.4	1356	4	014707	014707 homo sapien	958	6	3.7	32	2	09KLE0	09KLE0 vibrio chol
886	6	4.4	1364	13	090631	090631 gallus gall	959	6	3.7	32	10	09FU05	09FU05 zea mays su
887	6	4.4	1369	5	017176	017176 caenorhabdi	960	6	3.7	32	10	09FU03	09FU03 zea mays su
888	6	4.4	1369	11	063644	063644 rattus norv	961	6	3.7	32	10	09FU00	09FU00 zea luxuria
889	6	4.4	1435	5	09VTP9	09VTP9 drosophila	962	6	3.7	32	10	09FUP9	09FUP9 zea luxuria
890	6	4.4	1466	10	09ZUQ3	09ZUQ3 arabidopsis	963	6	3.7	32	10	09FEB2	09FEB2 zea mays su
891	6	4.4	1496	5	09NFE5	09NFE5 drosophila	964	6	3.7	32	10	09FEB1	09FEB1 zea mays su
892	6	4.4	1513	4	09NXY3	09NXY3 homo sapien	965	6	3.7	35	2	050052	050052 mycobacteri
893	6	4.4	1534	2	P71227	P71227 escherichia	966	6	3.7	37	4	013528	013528 homo sapien
894	6	4.4	1590	4	09P2N0	09P2N0 homo sapien	967	6	3.7	37	5	09NM39	09NM39 leishmania
895	6	4.4	1626	12	09YK98	09YK98 blackcurran	968	6	3.7	38	2	09RZM5	09RZM5 borrelia bu

```

969 5 3.7 38 5 018497 018497 caenorhabdi
970 5 3.7 39. 2 09KFE6 09f6c bacillus ha
971 5 3.7 39. 4 09Q013 09q13 homo sapien
972 5 3.7 41 5 076538 076538 strongyloce
973 5 3.7 42 6 062767 062767 equus cabal
974 5 3.7 44 11 004054 004054 rattus norv
975 5 3.7 45 11 09D289 09d289 mus musculu
976 5 3.7 45 11 09D289 09d289 mus musculu
977 5 3.7 45 11 09D289 09d289 mus musculu
978 5 3.7 45 13 09D052 09d052 bovine poly
979 5 3.7 48 5 08GR26 08gr26 aphidius co
980 5 3.7 50 12 09QZJ9 09qzj9 cavia porce
981 5 3.7 50 12 09J1K8 09j1k8 hepatitis c
982 5 3.7 52 12 096890 096890 hepatitis c
983 5 3.7 52 12 096890 096890 hepatitis c
984 5 3.7 53 5 09U5R8 09u5r8 entodinium
985 5 3.7 53 5 09N706 09n706 leishmania
986 5 3.7 54 9 09T0V8 09t0v8 bacterioph
987 5 3.7 54 12 055469 055469 human herpe
988 5 3.7 54 12 055470 055470 human herpe
989 5 3.7 56 2 099175 099175 porcine ent
990 5 3.7 56 2 09ZEV0 09zev0 bacillus st
991 5 3.7 56 2 09CFL7 09cfl7 lactococcus
992 5 3.7 56 11 09OZK0 09ozk0 cavia porce
993 5 3.7 56 12 056181 056181 human immun
994 5 3.7 56 12 056182 056182 human immun
995 5 3.7 56 12 056183 056183 human immun
996 5 3.7 56 12 056184 056184 human immun
997 5 3.7 56 12 056185 056185 human immun
998 5 3.7 56 12 056187 056187 human immun
999 5 3.7 58 2 0995U2 0995u2 staphylococ
1000 5 3.7 58 10 09LGD9 09lgd9 oryza sativ

```

## ALIGNMENTS

```

RESULT 1
099S07 PRELIMINARY: PRT: 163 AA.
AC 099S07;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE STAPHYLOKINASE PRECURSOR.
GN SAK OR SAI158.
OS Staphylococcus aureus subsp. aureus N315.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxId=158679;
RN [1]
RP SEQUENCE FROM N.A.
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Itoh T., Kanamori M.,
RA Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani U.I.,
RA Takehashi N.K., Sawano T., Inoue K.I., Kato C., Sekimizu K.,
RA Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,
RA Yanashita A., Osilima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
RA Ogasawara N., Hayashi H., Hirayatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancel 357:1225-1240(2001).
DR EMBL: AP003135, BAB43032.1; -.
KW Complete proteome.
SQ SEQUENCE 163 AA; 18520 MW; 413CDAE14BE9FD40 CRC64;

```

Query Match 100.0%; Score 136; DB 2; Length 163;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-132;  
 Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SSSFDKGRKYGDDASYFEPTGPLYLMVNTGVDSKGNELLSPHYVEPIKPGTTLTKRKI 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```

Db 28 SSSFDKGRKYGDDASYFEPTGPLYLMVNTGVDSKGNELLSPHYVEPIKPGTTLTKRKI 87
QY 61 EYVENALDATALKEFRVVELDPSAKIEVYTDKNKKKEETKSPITEKGFVVPDLSHI 120
Db 88 EYVENALDATALKEFRVVELDPSAKIEVYTDKNKKKEETKSPITEKGFVVPDLSHI 147
QY 121 KNGFNILTKVIEKK 136
Db 148 KNGFNILTKVIEKK 163

```

```

RESULT 2
033929 PRELIMINARY: PRT: 136 AA.
AC 033929;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 12, Last annotation update)
DE STAPHYLOKINASE (FRAGMENT).
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxId=1280;
RN [1]
RP SEQUENCE FROM N.A.
RA Chun H.S., Suk K., Kim S.H.;
RC STRAIN-ATCC 29213;
RA Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RL EMBL: U77328; AAB84174.1; -.
DR HSSP; P00802; ISSN.
FT NON_TER
SQ SEQUENCE 136 AA; 15551 MW; 79516BB136CA1A3F CRC64;

```

Query Match 73.5%; Score 100; DB 2; Length 136;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-95;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 37 NELLSPHYVEPIKPGTTLTKRKIEVYEMALDATALKEFRVVELDPSAKIEVYTDKNK 96
Db 37 NELLSPHYVEPIKPGTTLTKRKIEVYEMALDATALKEFRVVELDPSAKIEVYTDKNK 96
QY 97 KKEETKSPITEKGFVVPDLSHIKNGFNILTKVIEKK 136
Db 97 KKEETKSPITEKGFVVPDLSHIKNGFNILTKVIEKK 136

```

```

RESULT 3
09LC46 PRELIMINARY: PRT: 163 AA.
AC 09LC46;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE STAPHYLOKINASE.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxId=1280;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20219005; PubMed=10754251;
RA Horii T., Yokoyama K., Barua S., Odagiri T., Futamura N., Hasegawa T.,
RA Ohta M.;
RT "The staphylokinase gene is located in the structural gene encoding N-
RT acetylmuromyl-L-alanine amidase in methicillin-resistant
RT staphylococcus aureus."
RL FEMS Microbiol. Lett. 185:221-224(2000).
DR EMBL: AB033232; BAA93011.1; -.
KW Kinase.
SQ SEQUENCE 163 AA; 18474 MW; 3DBA5E35046029DD CRC64;

```

Query Match 71.3% Score 97; DB 2; Length 163;  
Best Local Similarity 100.0%; Pred. No. 3.9e-92;  
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 KGNELSPHYVEPIKPGTTLTKKEIEYVWALDATALYKFRVVELDPSAKIEVYDK 94  
|||||  
DB 62 KGNELSPHYVEPIKPGTTLTKKEIEYVWALDATALYKFRVVELDPSAKIEVYDK 121  
|||||

QY 95 NKKKEETKSPFITEKGVVDPDLSHKNPGNLITKV 131  
|||||  
DB 122 NKKKEETKSPFITEKGVVDPDLSHKNPGNLITKV 158  
|||||

RESULT 4  
Q9AM04 PRELIMINARY; PRT; 163 AA.

ID Q9AM04  
AC Q9AM04  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DE STAPHYLOKINASE SAKXH.  
OS STAPHYLOCOCCUS aureus.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Staphylococcus.  
OX NCBI\_TaxID=1280;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wei W., Xiang H., Tan H.;  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF332619; AAK11636.1; -  
KW Kinase.  
SQ SEQUENCE 163 AA; 18509 MW; 0A266B5BEAF65B5A CRC64;

Query Match 68.4% Score 93; DB 2; Length 163;  
Best Local Similarity 100.0%; Pred. No. 5.2e-88;  
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 YVEPIKPGTTLTKKEIEYVWALDATALYKFRVVELDPSAKIEVYDKNKKKEETKS 103  
|||||  
DB 71 YVEPIKPGTTLTKKEIEYVWALDATALYKFRVVELDPSAKIEVYDKNKKKEETKS 130  
|||||

QY 104 FPIITEKGVVDPDLSHKNPGNLITKVIEKK 136  
|||||  
DB 131 FPIITEKGVVDPDLSHKNPGNLITKVIEKK 163  
|||||

RESULT 5  
Q9X260 PRELIMINARY; PRT; 245 AA.

ID Q9X260  
AC Q9X260  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE GUPA PROTEIN.  
GN TM1738.  
OS Thermotoga maritima.  
OC Bacteria; Thermotogales; Thermotoga.  
OX NCBI\_TaxID=2336;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MSB8 / DSM 3109;  
RX MEDLINE=99287316; PubMed=10360571;  
RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,  
Hart D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,  
McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,  
Stewart A.M., Sutton M.D., Pratt M.S., Phillips C.A., Richardson D.,  
Heidelberg J., Smith H.O., Venter J.C., Fraser C.M.,  
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.,  
RT "Evidence for lateral gene transfer between Archaea and Bacteria from  
genome sequence of Thermotoga maritima.";  
RL Nature 399:323-329(1999).  
DR EMBL: AE001812; AAD36803.1; -

DR TIGR; TM1738; -  
DR InterPro; IPR000566; Lipocln\_cyFABP.  
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN\_1.  
KW Complete proteome.  
SQ SEQUENCE 245 AA; 25865 MW; 1FBAC67841ED2987 CRC64;

Query Match 5.9% Score 8; DB 2; Length 245;  
Best Local Similarity 100.0%; Pred. No. 4.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 SKGNELLS 41  
|||||  
DB 219 SKGNELLS 226  
|||||

RESULT 6  
Q27388 PRELIMINARY; PRT; 56 AA.

ID Q27388  
AC Q27388  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE HYPOTHETICAL 6.4 KDA PROTEIN.  
GN MTH1333.  
OS Methanobacterium thermoautotrophicum.  
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
OC Methanothermobacter.  
OX NCBI\_TaxID=145262;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DELTA H;  
RX MEDLINE=98037514; PubMed=9371463;  
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,  
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Olu D.,  
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
RA Ylvan N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
RA McQuigall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.;  
RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
delta: functional analysis and comparative genomics.";  
RL J. Bacteriol. 179:7135-7155(1997).  
DR EMBL: AE000897; AAB85811.1; -  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 56 AA; 6351 MW; 7709ECAC767AB635 CRC64;

Query Match 5.1% Score 7; DB 1; Length 56;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 TLTIREK 59  
|||||  
DB 16 TLTIREK 22  
|||||

RESULT 7  
Q28835 PRELIMINARY; PRT; 90 AA.

ID Q28835  
AC Q28835  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-NOV-1996 (TREMBlrel. 01, Last annotation update)  
DE CGMP-GATED RETINAL PHOTORECEPTOR CHANNEL (FRAGMENT).  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hundal S.P., DiFrancisco D., Mangoni M., Brammar W.J., Conley E.C.,  
RL Biochem. Soc. Trans. 21:0-0(1993).

DR EMBL: S65218; AAB27924.1; -  
 FT NON\_TER 1  
 SQ SEQUENCE 90 AA; 10981 MW; D6843EDBB87836C CRC64;

Query Match  
 Best Local Similarity 100.0%; Score 7; DB 6; Length 90;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 96 KKKEETK 102  
 |||||  
 Db 6 KKKEETK 12

RESULT 8  
 ID 086603 PRELIMINARY; PRT; 92 AA.  
 AC 086603;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE COAT PROTEIN (FRAGMENT).  
 OS Tulip top-breaking virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;  
 OC Potyviruses.  
 OX NCBI\_Taxid=32621;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93260402; PubMed=8492092;  
 RA Dekker E.L., Derks A.F., Asjes C.J., Lemmers M.E., Bol J.F.,  
 RA Langeveld S.A.;  
 RT "Characterization of potyviruses from tulip and lily which cause  
 RT flower-breaking.";  
 RL J. Gen. Virol. 74:881-887(1993).  
 DR EMBL: S60806; AAB2636.1;  
 DR InterPro: IPR001592; Pcty\_coat.  
 DR Pfam: PF00767; Pcty\_coat; 1.  
 KW Coat protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 92 AA; 10891 MW; 752DFCE2E0FC2B10 CRC64;

Query Match  
 Best Local Similarity 100.0%; Score 7; DB 12; Length 92;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 VEPPIKP 51  
 |||||  
 Db 18 VEPPIKP 24

RESULT 9  
 ID 09K8J5 PRELIMINARY; PRT; 102 AA.  
 AC 09K8J5;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE 50S RIBOSOMAL PROTEIN L21.  
 GN RPLU OR BH3011.  
 OS Bacillus halodurans.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 OX NCBI\_Taxid=6665;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C-125 / JCM 9153;  
 RX MEDLINE=20512582; PubMed=11058132;  
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
 RA Horikoshi K.;  
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";

RL Nucleic Acids Res. 28:4317-4331(2000).  
 CC -1- FUNCTION: THIS PROTEIN BINDS TO 23S RIBOSOMAL RNA IN THE PRESENCE  
 CC OF PROTEIN L20 (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE L21P FAMILY OF RIBOSOMAL PROTEINS.  
 DR EMBL: AP00517; BAB06730.1;  
 DR InterPro: IPR001787; Ribosomal\_L21P.  
 DR Pfam: PF00829; Ribosomal\_L21P; 1.  
 DR PROSITE: PS01169; Ribosomal\_L21; 1.  
 KW Complete proteome; Ribosomal protein; rRNA-binding.  
 SQ SEQUENCE 102 AA; 11350 MW; C929D2F40D75E3FC CRC64;

Query Match  
 Best Local Similarity 100.0%; Score 7; DB 2; Length 102;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 129 TKVIEK 135  
 |||||  
 Db 93 TKVIEK 99

RESULT 10  
 ID 09JEH2 PRELIMINARY; PRT; 114 AA.  
 AC 09JEH2;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DE REVERSE TRANSCRIPTASE (FRAGMENT).  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroviruses; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RT85;  
 RX MEDLINE=20148943; PubMed=10684266;  
 RA Goujon C.P., Schneider V.M., Groft J., Montigny J., Jeanfils V.,  
 RA Astagneau P., Rozenbaum W., Lot F., Frocraïn-Herchovitch C.,  
 RA Delphin N., Le Gal F., Nicolas J.-C., Milinkovich M.C., Deny P.;  
 RT "Phylogenetic Analyses indicate an Atypical Nurse-to-Patient  
 RT Transmission of Human Immunodeficiency Virus Type 1.";  
 RL J. Virol. 74:2525-2532(2000).  
 CC -1- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE  
 CC TRANSCRIPTASE).  
 DR EMBL: AF125749; AAF72270.1;  
 DR InterPro: IPR000477; RVase.  
 DR Pfam: PF00078; rvt; 1.  
 KW RNA-directed DNA polymerase.  
 FT NON\_TER 1  
 SQ SEQUENCE 114 AA; 13140 MW; 2D7D12027706430F CRC64;

Query Match  
 Best Local Similarity 100.0%; Score 7; DB 12; Length 114;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 55 LTKKIE 61  
 |||||  
 Db 2 LTKKIE 8

RESULT 11  
 ID 09KFP3 PRELIMINARY; PRT; 138 AA.  
 AC 09KFP3;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DE TRANSPOSASE (14).  
 GN BH0436.  
 OS Bacillus halodurans.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;



```

OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C-125 / JCM 9153;
RX MEDLINE=20512582; Pubmed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hiramata C., Nakamura Y., Ogasawara N., Kohara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL, AF001508; BAB04155.1; -.
KW Complete proteome.
SQ SEQUENCE 138 AA; 16118 MW; FF0BE8229CB5ED82 CRC64;

Query Match 5.1%; Score 7; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 DSKGNEL 39
   |||||
Db 80 DSKGNEL 86

RESULT 12
Q9QU35 PRELIMINARY; PRT; 140 AA.
AC Q9QU35;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE DNA, COMPLETE GENOME, ISOLATE:TLMV-CBD203.
OS TTV-like mini virus.
OC Viruses: ssDNA viruses; Circoviridae.
OX NCBI_TaxID=93678;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TLMV-CBD203;
RA Mishiro S.;
RT Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TLMV-CBD203;
RA Takahashi K., Iwasa Y., Hijikata M., Mishiro S.;
RT "Identification of a new human DNA virus (TTV-like mini virus: TLMV)
RT intermediately related to TTV virus and chicken anemia virus.";
RL Arch. Virol. 0:0-0(1999).
DR EMBL, AB026929; BAA86946.1; -.
SQ SEQUENCE 140 AA; 16573 MW; 607486347884FCE9 CRC64;

Query Match 5.1%; Score 7; DB 12; Length 140;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 TTTTKEK 59
   |||||
Db 60 TTTTKEK 66

RESULT 13
Q9WI01 PRELIMINARY; PRT; 154 AA.
AC Q9WI01;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE PROTEASE (FRAGMENT).
OS Human immunodeficiency virus type 2.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11709;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HEM-13;
RA Rodas B., Holguin A., Soriano V., Dourana M., Mansinho K., Antunes F.,
RA Gonzalez-Lahoz J.M.;
RT "Drug Resistance Mutations in HIV-2-infected Subjects Under
RT Antiretroviral Therapy: Identification of the 151 Multidrug-Resistant
RT Codon in Two Patients.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
DR EMBL: AF139053; AAD37335.1; -.
DR InterPro: IPR001969; Asp-Protease.
DR InterPro: IPR001995; Asp-Proteol.
DR Pfam: PF00077; rvp; 1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS0175; ASP_PROT_RETROV; 1.
KW Aspartyl protease; Hydrolase; Protease.
FT NON_TER 1
FT NON_TER 154
SQ SEQUENCE 154 AA; 16821 MW; C47EB4C660405791 CRC64;

Query Match 5.1%; Score 7; DB 12; Length 154;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 LTKERIE 61
   |||||
Db 124 LTKERIE 130

RESULT 14
Q9NF98 PRELIMINARY; PRT; 181 AA.
AC Q9NF98;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE HYPOTHETICAL 21.8 KDA PROTEIN.
GN PFC1016W.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=99376085; Pubmed=10448855;
RA Bowman S., Lawson D., Basham D., Brown D., Chillingworth T.,
RA Churcher C.M., Craig A., Davies R.M., Devlin K., Felwell T.,
RA Gentles S., Gilliam R., Hamlin N., Harris D., Holroyd S., Hornsby T.,
RA Horrocks P., Jagels K., Jassal B., Kyes S., McLean J., Moulé S.,
RA Mungall K., Murphy L., Oliver K., Quail M.A., Rajandream M.-A.,
RA Rutter S., Skelton J., Squares R., Squares S., Sulston J.E.,
RA Whitehead S., Woodward J.R., Newbold C., Barrall B.G.;
RT "The complete nucleotide sequence of chromosome 3 of Plasmodium
RT falciparum.";
RL Nature 400:532-538(1999).
DR EMBL: AL034559; CAB92305.1; -.
KW Hypothetical protein.
SQ SEQUENCE 181 AA; 21787 MW; 1F7CD3170DBD70AA CRC64;

Query Match 5.1%; Score 7; DB 5; Length 181;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 DKNKKE 99
   |||||
Db 6 DKNKKE 12

RESULT 15

```

Q9ACCO  
 ID Q9ACCO PRELIMINARY: PRT: 185 AA.  
 AC Q9ACCO:  
 DT 01-JUN-2001 (TREMBLER. 17, Created)  
 DT 01-JUN-2001 (TREMBLER. 17, Last sequence update)  
 DT 01-JUN-2001 (TREMBLER. 17, Last annotation update)  
 DE OUTER SURFACE PROTEIN PRECURSOR (FRAGMENT).  
 GN WSP.  
 OS Wolbachia endosymbiont of Heliella brevicornis.  
 OC Bacteria: Proteobacteria: alpha subdivision: Rickettsiales:  
 OC Rickettsiaceae; Wolbachiae; Wolbachia.  
 OX NCBI\_TaxId=118723;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA Cordaux R., Michel-Salzat A., Bouchon D.;  
 RT "Wolbachia infection in crustaceans: novel hosts and potential routes  
 RT for horizontal transmission";  
 RL J. Evol. Biol. 14:237-243(2001).  
 DR EMBL; AJ276609; CAC34465.1; -;  
 FT NON\_TER 1 185  
 FT NON\_TER 1 185  
 SQ SEQUENCE 185 AA; 19928 MW; 026198CA9EDFDE0C CRC64;

Query Match 5.1%; Score 7; DB 2; Length 185;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 YKKGDDA 15  
 |||||  
 Db 24 YKKGDDA 30

Search completed: April 22, 2002, 10:48:48  
 Job time: 203 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 10:38:39 ; Search time 23.68 seconds  
(without alignments)  
425.421 Million cell updates/sec

Title: US-09-601-490-1  
Perfect score: 715  
Sequence: 1 SSSFDKGYKGGDASYFEP.....SEHIKNPGFNLIKVVIEKK 136

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_1101.\*  
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	715	100.0	136	17 AAW03071	Wild-type staphylo
2	715	100.0	136	17 AAW03086	Wild-type staphylo
3	715	100.0	136	22 AAB99421	Staphylokinase (Sa
4	715	100.0	136	22 AAB61909	S. aureus staphylo
5	715	100.0	137	14 AAB39150	Staphylokinase SAK
6	712	99.6	136	20 AAY15024	Staphylokinase var
7	711	99.4	136	19 AAW44696	Staphylokinase fro
8	711	99.4	136	20 AAY15022	Staphylokinase var
9	711	99.4	136	20 AAY15025	Staphylokinase var
10	711	99.4	136	20 AAY15016	Staphylokinase var
11	711	99.4	136	20 AAY15021	Staphylokinase var

12	711	99.4	136	22 AAB61908	S. aureus staphylo
13	711	99.4	137	14 AAR39149	Staphylokinase SAK
14	711	99.4	163	12 AAR12137	S.aureus Staphylo
15	711	99.4	163	13 AAR28844	Staphylokinase (SAK
16	710	99.3	136	20 AAY15013	Staphylokinase var
17	709	99.2	136	17 AAW03101	Staphylokinase der
18	709	99.2	136	20 AAY15023	Staphylokinase var
19	709	99.2	136	20 AAY15015	Staphylokinase var
20	709	99.2	136	20 AAY15010	Staphylokinase var
21	709	99.2	136	20 AAY15004	Staphylokinase var
22	709	99.2	136	20 AAY15006	Staphylokinase var
23	708	99.0	136	19 AAW44690	Staphylokinase mut
24	708	99.0	136	19 AAW44693	Staphylokinase mut
25	708	99.0	136	20 AAY15026	Staphylokinase var
26	708	99.0	136	20 AAY15027	Staphylokinase var
27	707	98.9	136	19 AAW44694	Staphylokinase mut
28	707	98.9	136	20 AAY15007	Staphylokinase var
29	707	98.9	136	20 AAY15018	Staphylokinase var
30	707	98.9	136	20 AAY15019	Staphylokinase var
31	707	98.9	136	20 AAY15011	Staphylokinase var
32	706	98.7	136	19 AAW44689	Staphylokinase mut
33	706	98.7	136	19 AAW44692	Staphylokinase mut
34	706	98.7	136	20 AAY15028	Staphylokinase var
35	706	98.7	136	20 AAY15014	Staphylokinase var
36	706	98.7	136	21 AAB01297	Wild type staphylo
37	705	98.6	136	19 AAW44691	Staphylokinase mut
38	705	98.6	136	19 AAW44695	Staphylokinase mut
39	704	98.5	136	20 AAY15041	Staphylokinase var
40	704	98.5	136	20 AAY15020	Staphylokinase var
41	704	98.5	136	20 AAY15012	Staphylokinase var
42	703	98.3	136	17 AAW03079	Staphylokinase der
43	703	98.3	136	17 AAW03078	Staphylokinase der
44	703	98.3	136	17 AAW03082	Staphylokinase der
45	703	98.3	136	17 AAW03076	Staphylokinase der

ALIGNMENTS

RESULT 1  
AAW03071

ID AAW03071 standard; protein; 136 AA.

XX

AC AAW03071;

XX 19-FEB-1997 (first entry)

DT Wild-type staphylokinase.

XX

DE

XX

KW Staphylokinase; mutant; mutein; variant; immunogenicity; decrease;

KW derivative; SakSTAR; arterial thrombosis; thrombolytic agent.

XX

OS Staphylococcus aureus.

XX

XX Key Location/Qualifiers

FT Region 5..6

FT /label= M20\_epitope

FT Region 8..10

FT /label= M21\_epitope

FT Region 11..14

FT /label= M1\_epitope

FT Region 19

FT /label= M22\_epitope

FT Region 33..35

FT /label= M2\_epitope

FT Region 35..38

FT /label= M3\_epitope

FT Region 46..50

FT /label= M4\_epitope

FT Region 57..59

FT /label= M5\_epitope

FT Region 61..65

FT /label= M6\_epitope

*prior art*



PR 06-JAN-1995; 95EP-0200023.  
 PR 11-JAN-1995; 95US-0371505.  
 PR 09-JUN-1995; 95EP-0201531.  
 PR 06-JUL-1995; 95US-0499092.  
 XX  
 PA (COLL/) COLLEN D.  
 PA (LEUV-) LEUVEN RES & DEV VZW.  
 XX  
 XX Collen D;  
 XX  
 XX WPI; 1996-333991/33.

XX New staphylokinase derivs. having reduced immunogenicity - useful  
 PT for treating arterial thrombosis  
 XX

PS Disclosure; Fig 1; 58pp; English.

XX Staphylokinase derivatives showing a reduced immunogenicity as  
 CC compared to wild-type staphylokinase are claimed. The derivatives  
 CC are useful as thrombolytic agents to treat arterial thrombosis and  
 CC are pref. produced by eliminating at least one of the epitopes  
 CC indicated in the features table. The epitopes are destroyed by  
 CC replacing one or more amino acid residues in a charge cluster by an  
 CC Ala residue. Mutations are introduced using site-directed  
 CC mutagenesis on wild-type staphylokinase genomic DNA from the S.aureus  
 CC lysogenic strain SakSTAR. The present sequence is that of wild-type  
 CC SakSTAR staphylokinase.  
 XX  
 XX Sequence 136 AA;

Query Match 100.0%; Score 715; DB 17; Length 136;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-70;  
 Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSSFDKGYKKGDDASYFEPTGYLMVNVTVGDSKGNELLSPHYVEPIKPGTTLTKEKI 60  
 |||||  
 Db 1 sssfdkgkkgddasyfeptgpylmvntvgdskgnellspkyvefpikpgttltkeki 60  
 QY 61 EYVWEALDATALYKEFRVVELDPSAKIEVTYDYDNKKKKEETKSPFITEKGFVVPDLSEHI 120  
 |||||  
 Db 61 eyyvewaldataykefrvvelpsakievtydydnkkkeetksfitekgfvvpdlsehi 120  
 QY 121 KNPCFNLTITKVIEKK 136  
 |||||  
 Db 121 knpgfnlithkvviekk 136

RESULT 3

ID AAB99421  
 AC AAB99421; standard; Protein; 136 AA.

DT 28-AUG-2001 (first entry)

XX Staphylokinase (SakSTAR) primary protein sequence.

XX Staphylokinase; Staphylococcus aureus; SakSTAR; T-cell epitope;  
 KW immunogenic; thrombolytic; acute myocardial infarction; immunogenicity.  
 XX Staphylococcus aureus.

XX WO200140281-A2.

XX 07-JUN-2001.

XX 01-DEC-2000; 2000WO-EP12299.

XX 02-DEC-1999; 99EP-0204093.

XX (THRO-) THROMB-X NV.

XX

PI Warmerdam PAM, Plaisance SDNGH, Collen DJ, De Maeyer MCH;  
 XX  
 DR WPI; 2001-374786/39.

XX Reducing immunogenicity of protein (P) by eliminating T cell epitopes  
 PT in test peptides having amino acid sequence corresponding to (P) and  
 PT modifying amino acid sequence of (P) according to test peptide  
 PT modifications

XX Disclosure; Fig 1; 50pp; English.

XX The present invention describes a method for reducing the immunogenicity  
 CC of a peptide or protein (I), where (I) can be staphylokinase (SakSTAR)  
 CC isolated from Staphylococcus aureus. The method involves designing a  
 CC series of overlapping test peptides having an amino acid sequence  
 CC corresponding to (I), modifying test peptides which are identified to  
 CC comprise one or more T-cell epitopes, such that they are reduced or  
 CC eliminated, and repeating the T-cell eliminating modifications for (I)  
 CC to produce a modified peptide or protein. Methods from the present  
 CC invention can be used for the treatment, diagnosis or prophylaxis or  
 CC for the preparation of a pharmaceutical composition for the treatment,  
 CC diagnosis or prophylaxis of a human subject. Staphylokinase is used as a  
 CC potent thrombolytic agent in patients with acute myocardial infarction.  
 CC The method is useful for reducing cell based immunogenicity of non-human  
 CC proteins such as streptokinase or antibodies or their fragments, from  
 CC other species, for diagnostics and treatment of human disease. AAB99400  
 CC to AAB99449 represent amino acid sequences used in the exemplification  
 CC of the present invention.  
 XX  
 XX Sequence 136 AA;

Query Match 100.0%; Score 715; DB 22; Length 136;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-70;  
 Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSSFDKGYKKGDDASYFEPTGYLMVNVTVGDSKGNELLSPHYVEPIKPGTTLTKEKI 60  
 |||||  
 Db 1 sssfdkgkkgddasyfeptgpylmvntvgdskgnellspkyvefpikpgttltkeki 60  
 QY 61 EYVWEALDATALYKEFRVVELDPSAKIEVTYDYDNKKKKEETKSPFITEKGFVVPDLSEHI 120  
 |||||  
 Db 61 eyyvewaldataykefrvvelpsakievtydydnkkkeetksfitekgfvvpdlsehi 120  
 QY 121 KNPCFNLTITKVIEKK 136  
 |||||  
 Db 121 knpgfnlithkvviekk 136

RESULT 4

ID AAB61909  
 AC AAB61909 standard; Protein; 136 AA.

XX AAB61909;

DT 08-MAY-2001 (first entry)

XX S. aureus staphylokinase G34S variant.

XX Immunogenicity; staphylokinase; variant; stability.

XX Staphylococcus aureus.

XX Key Location/Qualifiers

FT Misc-difference 34

FT /label= G34S

FT /note= "wild-type Gly is replaced by Ser"

XX WO200104287-A1.

XX 18-JAN-2001.

XX 06-JUL-2000; 2000WO-DK00371.

XX

XX 07-JUL-1999; 99DK-0000988.  
PR 27-AUG-1999; 99DK-0001196.  
PR 02-MAR-2000; 2000DK-0000339.  
PR 18-MAY-2000; 2000DK-0000804.  
XX  
PA (MAXY-) MAXYGEN APS.  
XX  
PI Halkier T, Pedersen AH, Okkels JS;  
XX  
XX WPI; 2001-138342/14.  
DR  
XX  
XX Producing polypeptides with altered immunogenicity or improved  
PT stability, comprises expressing a diversified nucleotide sequence  
PT population and selecting polypeptides with altered immunogenicity or  
PT improved stability -  
XX  
PS  
PS Example 1; Page 76; 83pp; English.  
XX  
XX The invention relates to a method of altering immunogenicity and/or  
CC increasing stability of a polypeptide of interest. The method comprises  
CC (a) expressing a diversified population of nucleotide sequences encoding  
CC a polypeptide of interest; (b) screening the polypeptides expressed for  
CC function, immunogenicity and/or stability; and (c) selecting functional  
CC polypeptides with altered immunogenicity and/or increased stability. The  
CC method is used to improve the properties of polypeptides, in particular  
CC to alter the immunogenicity and/or increase the functional in vivo half-  
CC life of the polypeptide. The method uses a high throughput system that  
CC makes it possible to search several orders of magnitude more polypeptides  
CC than is possible by previously known approaches. This enhances the chance  
CC of finding the optimal variant from the many thousands of variants that  
CC may be produced. The present sequence represents S. aureus staphylokinase  
CC variant sequence, used to exemplify the method of the invention.  
XX  
SQ Sequence 136 AA;

Query Match 100.0%; Score 715; DB 22; Length 136;  
Best Local Similarity 100.0%; Pred. No. 1.3e-70;  
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSSFDKGYKKGGDASVEPTGPLYMNVNVTGVDKGNELLSPHYVEPIKPGTTLTKEKI 60  
Db 1 sssfdkgykkggddasyfeptgplymnnvtgvdskgnellshphyvefpikpgttltkeki 60  
QY 61 EYVVEWALDATAYKEFRVVELDPSAKIEVTYDKNKKKEETKSPITEKGFVVPDLSEHI 120  
Db 61 eyyvewaldataykefrvveldpsakievtyydknkkkeetksfitekfgfvvpdlsehi 120  
QY 121 KNPGFNLTIKVVIKK 136  
Db 121 knpgfnltikvviekk 136

RESULT 5  
AAR39150  
ID AAR39150 standard; Protein; 137 AA.  
XX  
AC AAR39150;  
XX  
DT 03-DEC-1993 (first entry)  
DE Staphylokinase SAK-STAR.  
XX  
KW Signal sequence, plasminogen activator; thrombosis; staphylokinase;  
KW SAK.  
XX  
OS Staphylococcus aureus strain 23.  
XX  
PN W09313209-A.  
XX  
PD 08-JUL-1993.  
XX

PF 28-DEC-1992; 92WO-EP02989.  
XX  
PR 30-DEC-1991; 91DE-4143279.  
PR 22-JUN-1992; 92DE-4220516.  
PR 01-DEC-1992; 92DE-4240801.  
XX  
PA (MEDA-) MEDAC GES KLINISCHE SPEZIALPRAEPARATE.  
XX  
XX Albrecht S, Behnke D, Guehrs K, Hartmann M, Schlott B;  
PI WPI; 1993-227325/28.  
XX  
DR N-PSDB; AAQ44270.  
XX  
XX Staphylo-kinase (SAK) sequences lacking sequences for signal  
PT peptide(s) - for prodn. of proteins used as plasminogen  
PT activators in thrombosis treatment, and monoclonal antibodies  
PT against SAK  
XX  
XX Claim 3; Fig 3; 99pp; German.  
XX  
XX DNA encoding SAK lacking the signal peptide, is expressed  
CC intracellularly. This avoids the problem of fast degradation of the  
CC polypeptides or destruction of the host when expressed into the  
CC medium or into the periplasm respectively. High expression is  
CC possible and the chemically induced overprodn. is easy to handle.  
CC Also, the prods. are homogeneous.  
CC SAK-polypeptide derivs. are plasminogen activators for the  
CC treatment of thrombosis.  
XX  
SQ Sequence 137 AA;

Query Match 100.0%; Score 715; DB 14; Length 137;  
Best Local Similarity 100.0%; Pred. No. 1.3e-70;  
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSSFDKGYKKGGDASVEPTGPLYMNVNVTGVDKGNELLSPHYVEPIKPGTTLTKEKI 60  
Db 2 sssfdkgykkggddasyfeptgplymnnvtgvdskgnellshphyvefpikpgttltkeki 61  
QY 61 EYVVEWALDATAYKEFRVVELDPSAKIEVTYDKNKKKEETKSPITEKGFVVPDLSEHI 120  
Db 62 eyyvewaldataykefrvveldpsakievtyydknkkkeetksfitekfgfvvpdlsehi 121  
QY 121 KNPGFNLTIKVVIKK 136  
Db 122 knpgfnltikvviekk 137

RESULT 6  
AAY15024  
ID AAY15024 standard; Protein; 136 AA.  
XX  
AC AAY15024;  
XX  
DT 03-NOV-1999 (first entry)  
DE Staphylokinase variant SakSTAR (V132L).  
XX  
KW Staphylokinase; variant; immunogenic; specificity; derivative; cysteine;  
KW SakSTAR (V132L); anti-thrombotic; fibrinolytic; cardiac; veterinary;  
KW Staphylococcus aureus; myocardial infarction; arterial thrombosis;  
KW monoclonal antibody; polyethylene glycol; PEG; plasma clearance.  
XX  
OS Staphylococcus aureus.  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
XX Misc-difference 132 /note= "Wild type Val is substituted by Leu"  
FT  
XX  
PN W09940198-A2.  
XX

PD		12-AUG-1999.	
XX			
XX	04-FEB-1999;	99WO-EP00748.	
XX			
PR	06-FEB-1998;	98EP-0200365.	
PR	04-FEB-1998;	98EP-0200323.	
XX			
PA	(COLL/) COLLEN D J.		
PA	(LEUV-) LEUVEN RES & DEV VZW.		
XX			
PI	Collen DJ:		
XX			
DR	WPI; 1999-508504/42.		
XX			
PT	Staphylokinase derivatives with reduced immunogenicity, used for,		
PT	e.g. treatment of arterial thrombosis		
XX			
PS	Claim 7; Page -; 10lpp; English.		
XX			
CC	The present sequence is a specifically claimed Staphylokinase SakSTAR		
CC	variant. This variant has one aminoacid that has been substituted		
CC	by another aminoacid that reduces the reactivity with monoclonal		
CC	antibodies and absorption of SakSTAR-specific antibodies from plasma		
CC	of patients treated with staphylokinase. The derivatives can also be		
CC	substituted with cysteine modified with PEG to maintain the specific		
CC	activity and significantly reduce the plasma clearance. They have		
CC	altered immunogenicity without markedly reducing the specific activity.		
CC	This sequence has anti-thrombotic, cardiant activity and a fibrinolytic		
CC	potency in human plasma. The new staphylokinase derivatives are used		
CC	for treatment of arterial thrombosis, especially myocardial infarction.		
CC	The compositions can be used in human or veterinary practice.		
CC	Note: The present sequence is not shown in the specification, but is		
CC	derived from the Staphylococcus aureus wild type staphylokinase sequence		
CC	given in figure 1.		
XX			
SQ	Sequence 136 AA;		
Query Match 99.6%; Score 712; DB 20; Length 136;			
Best Local Similarity 99.3%; Pred. No. 2.8e-70;			
Matches 135; Conservative 1; Mismatches 0; Indels 0; Gaps 0;			
Qy	1 SSSFDKGYKKGGDDASYEFTPGYLMVNVTVGVDSKGNELLSPHYVEFFIKPQTTLTKKI	60	
Dd	1 sssfdkgkykkgddasyfeptgpylmvnavtgvskgnellspyyvefipkpgttitkeki	60	
Qy	61 EYYEWALDAYKFRFVELDPDSAKIEVTYVDKNKKKEETKSFPITRKGFVVPLSEHI	120	
Dd	61 eyyewaldataykefrfveldpsakievttydknkkkeetksfpitekgfvvpdlsehi	120	
Qy	121 KNPGNLITKVIVIEKK 136		
Dd	121 knpgnlltkvliekx 136		
RESULT 7			
AAW44696			
ID	AAW44696 standard; protein; 136 AA.		
XX			
AC	AAW44696;		
XX			
DT	01-MAY-1998 (first entry)		
XX			
DE	Staphylokinase from Staphylococcus aureus.		
XX			
KW	Staphylokinase mutant; thrombolytic agent; Staphylococcus aureus.		
XX			
OS	Staphylococcus aureus.		
XX			
PN	JPI0028587-A.		
XX			
PD	03-FEB-1998.		

PR 04-FEB-1998; 98EP-0200323.  
 XX (COLL/) COLLEN D J.  
 PA (LEUV-) LEUVEN RES & DEV VZW.  
 XX Collen DJ;  
 PI  
 XX WPI; 1999-508504/42.  
 DR Staphylokinase derivatives with reduced immunogenicity, used for,  
 XX e.g. treatment of arterial thrombosis  
 PT Claim 7; Page -: 101pp; English.  
 XX  
 CC The present sequence is a specifically claimed Staphylokinase SakSTAR  
 CC variant. This variant has one aminoacid that has been substituted  
 CC by another aminoacid that reduces the reactivity with monoclonal  
 CC antibodies and absorption of SakSTAR-specific antibodies from plasma  
 CC of patients treated with staphylokinase. The derivatives can also be  
 CC substituted with cysteine modified with PEG to maintain the specific  
 CC activity and significantly reduce the plasma clearance. They have  
 CC altered immunogenicity without markedly reducing the specific activity.  
 CC This sequence has anti-thrombotic, cardiant activity and a fibrinolytic  
 CC potency in human plasma. The new staphylokinase derivatives are used  
 CC for treatment of arterial thrombosis, especially myocardial infarction.  
 CC The compositions can be used in human or veterinary practice.  
 CC Note: The present sequence is not shown in the specification, but is  
 CC derived from the Staphylococcus aureus wild type staphylokinase sequence  
 CC given in figure 1.  
 XX  
 SQ Sequence 136 AA;

Query Match 99.4%; Score 711; DB 20; Length 136;  
 Best Local Similarity 99.3%; Pred. No. 3.6e-70;  
 Matches 135; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SSSFDKGYKKGDDASYFEPTGPLYMNVTVGDSKGNELLSPHYVEFPKPGTTLTKEKI 60  
 DB 1 sssfdkgykkgddasyfeptgplymnvntgvdskgnellsphyvefpikpgtlttkeki 60  
 QY 61 EYVVEWALDATAYKEFRVVELDPSAKIEVTYYDKNKKKEETKSPFITEKGFVVVPDLSEHI 120  
 DB 61 eyvvewaldataykefrvveldpsakievtyydknkkkeetksfpitekgyfvvpdlsehi 120  
 QY 121 KNPGFNLTIKVVEKK 136  
 DB 121 knpgfnltikvviekk 136

RESULT 9  
 AAY15025  
 ID AAY15025 standard; Protein; 136 AA.  
 XX  
 AC AAY15025;  
 XX  
 DT 03-NOV-1999 (first entry)  
 XX  
 DE Staphylokinase variant SakSTAR (V132T).  
 XX  
 KW Staphylokinase; variant; immunogenic; specificity; derivative; cysteine;  
 KW SakSTAR (V132T); anti-thrombotic; fibrinolytic; cardiant; veterinary;  
 KW Staphylococcus aureus; myocardial infarction; arterial thrombosis;  
 KW monoclonal antibody; polyethylene glycol; PEG; plasma clearance.  
 XX  
 OS Staphylococcus aureus.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 132  
 FT /note= "wild type val is substituted by Thr"  
 XX  
 PN W09940198-A2.

XX 12-AUG-1999.  
 PD  
 XX  
 PF 04-FEB-1999; 99WO-EP00748.  
 XX  
 PR 06-FEB-1998; 98EP-0200365.  
 PR 04-FEB-1998; 98EP-0200323.  
 XX  
 XX (COLL/) COLLEN D J.  
 PA (LEUV-) LEUVEN RES & DEV VZW.  
 XX Collen DJ;  
 PI  
 XX WPI; 1999-508504/42.  
 DR Staphylokinase derivatives with reduced immunogenicity, used for,  
 XX e.g. treatment of arterial thrombosis  
 PT Claim 7; Page -: 101pp; English.  
 XX

CC The present sequence is a specifically claimed Staphylokinase SakSTAR  
 CC variant. This variant has one aminoacid that has been substituted  
 CC by another aminoacid that reduces the reactivity with monoclonal  
 CC antibodies and absorption of SakSTAR-specific antibodies from plasma  
 CC of patients treated with staphylokinase. The derivatives can also be  
 CC substituted with cysteine modified with PEG to maintain the specific  
 CC activity and significantly reduce the plasma clearance. They have  
 CC altered immunogenicity without markedly reducing the specific activity.  
 CC This sequence has anti-thrombotic, cardiant activity and a fibrinolytic  
 CC potency in human plasma. The new staphylokinase derivatives are used  
 CC for treatment of arterial thrombosis, especially myocardial infarction.  
 CC The compositions can be used in human or veterinary practice.  
 CC Note: The present sequence is not shown in the specification, but is  
 CC derived from the Staphylococcus aureus wild type staphylokinase sequence  
 CC given in figure 1.  
 XX  
 SQ Sequence. 136 AA;

Query Match 99.4%; Score 711; DB 20; Length 136;  
 Best Local Similarity 99.3%; Pred. No. 3.6e-70;  
 Matches 135; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 SSSFDKGYKKGDDASYFEPTGPLYMNVTVGDSKGNELLSPHYVEFPKPGTTLTKEKI 60  
 DB 1 sssfdkgykkgddasyfeptgplymnvntgvdskgnellsphyvefpikpgtlttkeki 60  
 QY 61 EYVVEWALDATAYKEFRVVELDPSAKIEVTYYDKNKKKEETKSPFITEKGFVVVPDLSEHI 120  
 DB 61 eyvvewaldataykefrvveldpsakievtyydknkkkeetksfpitekgyfvvpdlsehi 120  
 QY 121 KNPGFNLTIKVVEKK 136  
 DB 121 knpgfnltikvviekk 136

RESULT 10  
 AAY15016  
 ID AAY15016 standard; Protein; 136 AA.  
 XX  
 AC AAY15016;  
 XX  
 DT 03-NOV-1999 (first entry)  
 XX  
 DE Staphylokinase variant SakSTAR (V132A).  
 XX  
 KW Staphylokinase; variant; immunogenic; specificity; derivative; cysteine;  
 KW SakSTAR (V132A); anti-thrombotic; fibrinolytic; cardiant; veterinary;  
 KW Staphylococcus aureus; myocardial infarction; arterial thrombosis;  
 KW monoclonal antibody; polyethylene glycol; PEG; plasma clearance.  
 XX  
 OS Staphylococcus aureus.  
 OS Synthetic.



XX Key Location/Qualifiers  
 FH Misc-difference 132  
 FT /note= "Wild type Val is substituted by Ala"  
 XX  
 XX WO9940198-A2.  
 XX  
 XX 12-AUG-1999.  
 XX  
 XX 04-FEB-1999; 99WO-EF00748.  
 XX  
 XX 06-FEB-1998; 98EP-0200365.  
 XX 04-FEB-1998; 98EP-0200323.  
 XX  
 XX (COLL/) COLLEN D J.  
 XX (LEUV-) LEUVEN RES & DEV VZW.  
 XX  
 XX Collen DJ;  
 XX  
 XX WPI; 1999-508504/42.  
 XX  
 XX Staphylokinase derivatives with reduced immunogenicity, used for,  
 XX e.g. treatment of arterial thrombosis  
 XX  
 XX Claim 7; Page -: 101pp; English.  
 XX  
 XX The present sequence is a specifically claimed Staphylokinase SakSTAR  
 CC variant. This variant has one aminoacid that has been substituted  
 CC by another aminoacid that reduces the reactivity with monoclonal  
 CC antibodies and absorption of SakSTAR-specific antibodies from plasma  
 CC of patients treated with staphylokinase. The derivatives can also be  
 CC substituted with cysteine modified with PEG to maintain the specific  
 CC activity and significantly reduce the plasma clearance. They have  
 CC altered immunogenicity without markedly reducing the specific activity.  
 CC This sequence has anti-thrombotic, cardiant activity and a fibrinolytic  
 CC potency in human plasma. The new staphylokinase derivatives are used  
 CC for treatment of arterial thrombosis, especially myocardial infarction.  
 CC The compositions can be used in human or veterinary practice.  
 CC Note: The present sequence is not shown in the specification, but is  
 CC derived from the Staphylococcus aureus wild type staphylokinase sequence  
 CC given in figure 1.  
 XX  
 XX Sequence 136 AA;  
 XX  
 Query Match 99.4%; Score 711; DB 20; Length 136;  
 Best Local Similarity 99.3%; Pred. No. 3.6e-70;  
 Matches 135; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 SSSFDKGYKKGDDASYFEPTGPLYMNVNVTGVDKGNELLSPHYVEPIKPGTTLTKEKI 60  
 Db 1 sssfdkgykkgddasyfeptgplymnnvtgvdskgnellspkpgttltkeki 60  
 QY 61 EYVVEWALDATAYKEFRVVELDPSAKIEVTYVDKNKKKEETKSPITEKGFVVPDLSEHI 120  
 Db 61 eyyvewaldataykefrvvelpsakievtyydknkkkeetksfpitkfgfvvpdlsehi 120  
 QY 121 KNPGFNLITKVIIEKK 136  
 Db 121 knpgfnlitrkviiekk 136  
 RESULT 11  
 ID AAY15021  
 XX AAY15021 standard; Protein; 136 AA.  
 XX  
 XX AAY15021;  
 XX  
 XX 03-NOV-1999 (first entry)  
 XX  
 XX Staphylokinase variant SakSTAR (K35E).  
 DE  
 XX Staphylokinase; variant; immunogenic; specificity; derivative: cysteine;  
 KW

KW SakSTAR (K35E); anti-thrombotic; fibrinolytic; cardiant; veterinary;  
 KW Staphylococcus aureus; myocardial infarction; arterial thrombosis;  
 KW monoclonal antibody; polyethylene glycol; PEG; plasma clearance.  
 XX  
 XX Staphylococcus aureus.  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 35  
 FT /note= "Wild type Lys is substituted by Glu"  
 XX  
 XX WO9940198-A2.  
 XX  
 XX 12-AUG-1999.  
 XX  
 XX 04-FEB-1999; 99WO-EF00748.  
 XX  
 XX 06-FEB-1998; 98EP-0200365.  
 XX 04-FEB-1998; 98EP-0200323.  
 XX  
 XX (COLL/) COLLEN D J.  
 XX (LEUV-) LEUVEN RES & DEV VZW.  
 XX  
 XX Collen DJ;  
 XX  
 XX WPI; 1999-508504/42.  
 XX  
 XX Staphylokinase derivatives with reduced immunogenicity, used for,  
 XX e.g. treatment of arterial thrombosis  
 XX  
 XX Claim 7; Page -: 101pp; English.  
 XX  
 XX The present sequence is a specifically claimed Staphylokinase SakSTAR  
 CC variant. This variant has one aminoacid that has been substituted  
 CC by another aminoacid that reduces the reactivity with monoclonal  
 CC antibodies and absorption of SakSTAR-specific antibodies from plasma  
 CC of patients treated with staphylokinase. The derivatives can also be  
 CC substituted with cysteine modified with PEG to maintain the specific  
 CC activity and significantly reduce the plasma clearance. They have  
 CC altered immunogenicity without markedly reducing the specific activity.  
 CC This sequence has anti-thrombotic, cardiant activity and a fibrinolytic  
 CC potency in human plasma. The new staphylokinase derivatives are used  
 CC for treatment of arterial thrombosis, especially myocardial infarction.  
 CC The compositions can be used in human or veterinary practice.  
 CC Note: The present sequence is not shown in the specification, but is  
 CC derived from the Staphylococcus aureus wild type staphylokinase sequence  
 CC given in figure 1.  
 XX  
 XX Sequence 136 AA;  
 XX  
 Query Match 99.4%; Score 711; DB 20; Length 136;  
 Best Local Similarity 99.3%; Pred. No. 3.6e-70;  
 Matches 135; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SSSFDKGYKKGDDASYFEPTGPLYMNVNVTGVDKGNELLSPHYVEPIKPGTTLTKEKI 60  
 Db 1 sssfdkgykkgddasyfeptgplymnnvtgvdskgnellspkpgttltkeki 60  
 QY 61 EYVVEWALDATAYKEFRVVELDPSAKIEVTYVDKNKKKEETKSPITEKGFVVPDLSEHI 120  
 Db 61 eyyvewaldataykefrvvelpsakievtyydknkkkeetksfpitkfgfvvpdlsehi 120  
 QY 121 KNPGFNLITKVIIEKK 136  
 Db 121 knpgfnlitrkviiekk 136  
 RESULT 12  
 ID AAB61908  
 XX AAB61908 standard; Protein; 136 AA.  
 XX  
 XX AAB61908;  
 AC

AC AAR39149;  
XX 03-DEC-1993 (first entry)  
XX Staphylokinase SAK-Cphic.  
DE Signal sequence, plasminogen activator; thrombosis; staphylokinase;  
KW SAK.  
XX Staphylococcus aureus phage phiC.  
OS WO9313209-A.  
XX 08-JUL-1993.  
XX 28-DEC-1992; 92WO-EP02989.  
XX 30-DEC-1991; 91DE-4143279.  
XX 22-JUN-1992; 92DE-4220516.  
XX 01-DEC-1992; 92DE-4240801.  
XX (MEDA-) MEDAC GES KLINISCHE SPEZIALPRAEPARATE.  
XX Albrecht S, Behnke D, Guehrs K, Hartmann M, Schlott B;  
XX WPI; 1993-227325/28.  
XX N-PSDB; AAQ44269.  
XX Staphylo-kinase (SAK) sequences lacking sequences for signal  
XX peptide(s) - for prodn. of proteins used as plasminogen  
XX activators in thrombosis treatment, and monoclonal antibodies  
XX against SAK  
XX Claim 3; Fig 2; 99pp; German.  
XX DNA encoding SAK lacking the signal peptide, is expressed  
XX intracellularly. This avoids the problem of fast degradation of the  
XX polypeptides or destruction of the host when expressed into the  
XX medium or into the periplasm respectively. High expression is  
XX possible and the chemically induced overprodn. is easy to handle.  
XX Also, the prods. are homogeneous. plasminogen activators for the  
XX SAK-polypeptide derivs. are plasminogen activators for the  
XX treatment of thrombosis.  
XX Sequence 137 AA;  
XX Query Match 99.4%; Score 711; DB 14; Length 137;  
XX Best Local Similarity 99.3%; Pred. No. 3.6e-70;  
XX Matches 135; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 SSSFDKGYKKGDDASYFEPTGPLYMVNVTGVDSKGNELLSPHYVEFPKPGTTLTKEKI 60  
Db 2 SSSfdkgykkgddasyfeptgplymvnvtgvdgkgnellspkpgtlttkeki 61  
QY 61 EYVVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSFPTTEKGFVVPDLSEHI 120  
Db 62 EYVVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSFPTTEKGFVVPDLSEHI 121  
QY 121 KNPGFNLITKVVEIEKK 136  
Db 122 knpgfnlntkvvlekk 137  
RESULT 14  
AAR12137  
ID AAR12137 standard; Protein; 163 AA.  
XX AAR12137;  
XX 05-AUG-1991 (first entry)  
XX S. aureus Staphylokinase with signal peptide.

AC AAR39149;  
XX 08-MAY-2001 (first entry)  
XX S. aureus staphylokinase mature protein sequence.  
DE Immunogenicity; staphylokinase; variant; stability.  
KW Staphylococcus aureus.  
OS WO200104287-A1.  
XX 18-JAN-2001.  
XX 06-JUL-2000; 2000WO-DK00371.  
XX 07-JUL-1999; 99DK-0000988.  
XX 27-AUG-1999; 99DK-0001196.  
XX 02-MAR-2000; 2000DK-0000339.  
XX 18-MAY-2000; 2000DK-0000804.  
XX (MAXY-) MAXYGEN APS.  
XX Halkier T, Pedersen AH, Okkels JS;  
XX WPI; 2001-138342/14.  
XX N-PSDB; AAC85112.  
XX Producing polypeptides with altered immunogenicity or improved  
XX stability, comprises expressing a diversified nucleotide sequence  
XX population and selecting polypeptides with altered immunogenicity or  
XX improved stability -  
XX Example 1; Page 75; 83pp; English.  
XX The invention relates to a method of altering immunogenicity and/or  
XX increasing stability of a polypeptide of interest. The method comprises  
XX (a) expressing a diversified population of nucleotide sequences encoding  
XX a polypeptide of interest; (b) screening the polypeptides expressed for  
XX function, immunogenicity and/or stability; and (c) selecting functional  
XX polypeptides with altered immunogenicity and/or increased stability. The  
XX method is used to improve the properties of polypeptides, in particular  
XX to alter the immunogenicity and/or increase the functional in vivo half-  
XX life of the polypeptide. The method uses a high throughput system that  
XX makes it possible to search several orders of magnitude more polypeptides  
XX than is possible by previously known approaches. This enhances the chance  
XX of finding the optimal variant from the many thousands of variants that  
XX may be produced. The present sequence represents S. aureus staphylokinase  
XX mature protein sequence, used to exemplify the method of the invention.  
XX Sequence 136 AA;  
XX Query Match 99.4%; Score 711; DB 22; Length 136;  
XX Best Local Similarity 99.3%; Pred. No. 3.6e-70;  
XX Matches 135; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 SSSFDKGYKKGDDASYFEPTGPLYMVNVTGVDSKGNELLSPHYVEFPKPGTTLTKEKI 60  
Db 1 SSSfdkgykkgddasyfeptgplymvnvtgvdgkgnellspkpgtlttkeki 60  
QY 61 EYVVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSFPTTEKGFVVPDLSEHI 120  
Db 61 EYVVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSFPTTEKGFVVPDLSEHI 120  
QY 121 KNPGFNLITKVVEIEKK 136  
Db 121 knpgfnlntkvvlekk 136  
RESULT 13  
AAR39149  
ID AAR39149 standard; Protein; 137 AA.  
XX

XX SAK; protein production.  
XX Synthetic.  
XX Key Location/Qualifiers  
XX Peptide 1..27  
XX Protein /label= signal peptide  
XX 28..163  
XX /label= staphylokinase  
XX JP03098595-A.  
XX PN  
XX PD 24-APR-1991.  
XX PF 11-SEP-1989; 89JP-0234874.  
XX PR 11-SEP-1989; 89JP-0234874.  
XX (TAIS ) TAISHO PHARMACEUT KK.  
XX WPI; 1991-167039/23.  
XX DR N-PSDB; AAQ11813.  
XX PT Prepn. of peptide(s) - by construction of expression  
XX vector, transformation of E.coli etc., culturing to secrete  
XX peptide(s) and collecting peptide(s)  
XX PS Example; Fig 3; 15pp; Japanese.  
XX CC The first 79 residues of this sequence form part of a fusion  
XX protein with somatomedin C. A synthetic construct encoding the  
XX SAK-SMC fusion is used to transform E.coli or Bacillus subtilis.  
XX CC The SAK signal peptide directs extracellular secretion of the  
XX somatomedin C.  
XX CC See also AAQ11814.  
XX SQ Sequence 163 AA;

Query Match 99.4%; Score 711; DB 12; Length 163;  
Best Local Similarity 99.3%; Pred. No. 4.6e-70;  
Matches 135; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 SSFDDKGYKKGDDASVFETGPLYMNVTCVDSKGNELLSPHYVEPIKPGTTLTKEKI 60  
DB 28 sssfdkgkykkgddasyfeptgplymnnvtgvdgkgnellsphyvefpikpgttltkeki 87  
QY 61 EYVVEWALDATAYKEFRVVELDPSAKIEVTYDKNKKKEETKSPITEKGFVVPDLSEHI 120  
DB 88 eyyvewaldataykefrvvelpsakievtyydknkkkeetksfpitekgfvvpdlsehi 147  
QY 121 KNPGFNLITKVVIEKK 136  
DB 148 knpgfnlitrkvviekk 163

RESULT 15  
ID AAR28844  
XX AAR28844 standard; Protein; 163 AA.  
XX AC AAR28844;  
XX 13-JAN-1993 (first entry)  
XX DE Staphylokinas (SAK).  
XX KW Staphylokinase; thrombosis; fibrinogen; plasminogen.  
XX OS Synthetic.  
XX FH Key Location/Qualifiers  
XX Peptide 1..27

FT Protein /label= sig\_peptide  
FT 28..163  
FT /label= mat\_protein  
XX WO9211356-A.  
XX PN  
XX PD 09-JUL-1992.  
XX PF 17-DEC-1991; 91WO-JP01722.  
XX PR 17-DEC-1990; 90JP-0411063.  
XX (HONS ) YAKULT HONSHA KK.  
XX Hashimoto S, Matsumoto T, Matsuo O, Onoue M, Sakai M;  
XX Sako T, Sansawa H, Shimura K, Shishido Y, Watanabe T;  
XX Yokokura T;  
XX WPI; 1992-250079/30.  
XX Thrombolytic peptide SAK-11 derived from staphylokinase - useful  
XX as a plasminogen activator for treating thrombosis  
XX PS Disclosure; Page 16; 26pp; Japanese.  
XX CC SAK-11 (AAR25468) is derived from its precursor, staphylokinase (SAK)  
XX - represented in AAR28844, by tryptic cleavage of its 10 N-terminal  
XX amino acid residues. SAK-11 is useful as the active agent in an  
XX injectable treatment for thrombosis. Thrombolytic activity is  
XX demonstrated using a rabbit jugular vein thrombosis model. A clear  
XX decrease in fibrinogen is observed upon application of 0.9 mg SAK-11/  
XX kg. Plasminogen activation is superior to that of SAK (no specific  
XX data given). Furthermore, SAK-11 has very low antigenicity as shown  
XX in a PCA study, against streptokinase and egg albumin controls, in  
XX BALB/c mice.  
XX SQ Sequence 163 AA;

Query Match 99.4%; Score 711; DB 13; Length 163;  
Best Local Similarity 99.3%; Pred. No. 4.6e-70;  
Matches 135; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 SSFDDKGYKKGDDASVFETGPLYMNVTCVDSKGNELLSPHYVEPIKPGTTLTKEKI 60  
DB 28 sssfdkgkykkgddasyfeptgplymnnvtgvdgkgnellsphyvefpikpgttltkeki 87  
QY 61 EYVVEWALDATAYKEFRVVELDPSAKIEVTYDKNKKKEETKSPITEKGFVVPDLSEHI 120  
DB 88 eyyvewaldataykefrvvelpsakievtyydknkkkeetksfpitekgfvvpdlsehi 147  
QY 121 KNPGFNLITKVVIEKK 136  
DB 148 knpgfnlitrkvviekk 163

Search completed: April 22, 2002, 10:41:30  
Job time: 171 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 10:40:19 ; Search time 12.47 Seconds  
(without alignments)  
245.425 Million cell updates/sec

Title: US-09-601-490-1  
Perfect score: 715  
Sequence: 1 SSSFDGKGYKGGDASYFEP.....SEHIKNPGFNLTIKVIEKK 136

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents.AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	715	100.0	136	1 US-08-371-505-2	Sequence 2, Appl
2	715	100.0	136	2 US-08-784-971-5	Sequence 5, Appl
3	715	100.0	137	1 US-08-256-261-6	Sequence 6, Appl
4	715	100.0	137	3 US-08-852-299-6	Sequence 6, Appl
5	711	99.4	137	1 US-08-256-261-4	Sequence 4, Appl
6	711	99.4	137	3 US-08-852-299-4	Sequence 4, Appl
7	705	98.6	163	1 US-08-075-545-1	Sequence 1, Appl
8	695	97.2	137	1 US-08-256-261-2	Sequence 2, Appl
9	695	97.2	137	3 US-08-852-299-2	Sequence 2, Appl
10	695	97.2	163	1 US-08-256-261-17	Sequence 17, Appl
11	695	97.2	163	3 US-08-852-299-17	Sequence 17, Appl
12	692	96.8	137	1 US-08-256-261-14	Sequence 14, Appl
13	692	96.8	137	3 US-08-852-299-14	Sequence 14, Appl
14	689	96.4	137	1 US-08-256-261-12	Sequence 12, Appl
15	689	96.4	137	3 US-08-852-299-12	Sequence 12, Appl
16	653	91.3	126	1 US-08-075-545-2	Sequence 2, Appl
17	643	89.9	127	1 US-08-256-261-8	Sequence 8, Appl
18	643	89.9	127	3 US-08-852-299-8	Sequence 8, Appl
19	620	86.7	123	1 US-08-256-261-10	Sequence 10, Appl
20	620	86.7	123	3 US-08-852-299-10	Sequence 10, Appl
21	74	10.3	592	4 US-09-000-145-6	Sequence 6, Appl
22	73.5	10.3	2522	4 US-09-251-645-13	Sequence 13, Appl
23	73	10.2	14	2 US-08-934-222-23	Sequence 23, Appl
24	73	10.2	14	2 US-08-933-402-23	Sequence 23, Appl
25	73	10.2	14	2 US-09-207-621-23	Sequence 23, Appl
26	73	10.2	14	2 US-08-532-818-23	Sequence 23, Appl
27	73	10.2	14	3 US-09-231-797-23	Sequence 23, Appl

28 73 10.2 14 3 US-08-934-224-23 Sequence 23, Appl  
29 73 10.2 14 3 US-08-933-843-23 Sequence 23, Appl  
30 73 10.2 14 4 US-08-934-223-23 Sequence 23, Appl  
31 73 10.2 14 4 US-08-933-843-23 Sequence 23, Appl  
32 73 10.2 770 1 US-08-525-654A-1 Sequence 1, Appl  
33 73 10.2 771 1 US-08-525-654A-3 Sequence 3, Appl  
34 73 10.2 2509 1 US-08-469-005A-10 Sequence 10, Appl  
35 73 10.2 2511 4 US-09-261-907-2 Sequence 2, Appl  
36 70.5 9.9 706 1 US-08-339-152A-29 Sequence 29, Appl  
37 70.5 9.9 706 2 US-08-007-999B-4 Sequence 4, Appl  
38 70.5 9.9 706 2 US-08-689-276A-4 Sequence 4, Appl  
39 69 9.7 423 1 US-08-476-008-8 Sequence 8, Appl  
40 69 9.7 423 1 US-08-306-063-8 Sequence 8, Appl  
41 69 9.7 423 1 US-08-833-485-8 Sequence 8, Appl  
42 69 9.7 423 4 US-09-137-440-8 Sequence 8, Appl  
43 69 9.7 423 5 PCT-US91-06148A-8 Sequence 8, Appl  
44 69 9.7 427 4 US-09-243-374-7 Sequence 7, Appl  
45 69 9.7 427 6 5310667-7 Patent No. 5310667

ALIGNMENTS

RESULT 1  
US-08-371-505-2  
; Sequence 2, Application US/08371505  
; Patent No. 5695754  
; GENERAL INFORMATION:  
; APPLICANT: COLLEN, DESIRE  
; TITLE OF INVENTION: STAPHYLOKINASE DERIVATIVES  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WEBB ZIESENHEIM BRUENING LOGSDON ORKIN & HANSON, P.C.  
; STREET: 700 KOPPERS BUILDING, 436 SEVENTH AVENUE  
; CITY: PITTSBURGH  
; STATE: PENNSYLVANIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 15222-2363  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: NEC 286  
; OPERATING SYSTEM: DOS  
; SOFTWARE: WORDPERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/371,505  
; FILING DATE: 11 JAN 1995  
; CLASSIFICATION: 424  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 136  
; TYPE: AMINO ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: UNKNOWN  
US-08-371-505-2

Query Match 100.0%; Score 715; DB 1; Length 136;  
Best Local Similarity 100.0%; Pred. No. 1.4e-78;  
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSSFDGKGYKGGDASYFEP...SEHIKNPGFNLTIKVIEKK 60  
Db 1 SSSFDGKGYKGGDASYFEP...SEHIKNPGFNLTIKVIEKK 60  
QY 61 EYVEWALDATAATAYKEFRVVELDPSAKIEV...SEHIKNPGFNLTIKVIEKK 120  
Db 61 EYVEWALDATAATAYKEFRVVELDPSAKIEV...SEHIKNPGFNLTIKVIEKK 120  
QY 121 KNPGFNLTIKVIEKK 136  
Db 121 KNPGFNLTIKVIEKK 136



US-08-852-299-6

	Query Match	100.0%;	Score 715;	DB 3;	Length 137;
	Best Local Similarity	100.0%;	Pred. No. 1.4e-78;		
	Matches 136;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	SSSPDKGKYKGGDASFEPTGPYLMWNTGVDGSGNELLSPHYVEPIKPGTTLTKRKI	60		
Db	2	SSSPDKGKYKGGDASFEPTGPYLMWNTGVDGSGNELLSPHYVEPIKPGTTLTKRKI	61		
Qy	61	EYYVEWALDATAYKEFRVVELDPSAKIEVTYYDKNKKKEETKSPITEKGVVPDLSHI	120		
Db	62	EYYVEWALDATAYKEFRVVELDPSAKIEVTYYDKNKKKEETKSPITEKGVVPDLSHI	121		
Qy	121	KNPGFNLTIKVVIIEKK	136		
Db	122	KNPGFNLTIKVVIIEKK	137		

```

RESULT      5
US-08-256-261-4
; Sequence 4, Application US/08256261
; Patent No. 5801037
; GENERAL INFORMATION:
; APPLICANT: Behnke, Detlef
; APPLICANT: Schlott, Bernhard
; APPLICANT: Albrecht, Sybille
; APPLICANT: G hrs, Karl-Helinz
; APPLICANT: Hartmann, Manfred
; TITLE OF INVENTION: Expression of signal-peptide-free
;                   staphylokinases
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:

```

```

;
; COMPUTER: IBM PC COMPACTIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,261
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 137 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; MS-08-256-261-4

```

Query Match	99.4%	Score 711;	DB 1;	Length 137;
Best Local Similarity	99.3%;	Pred. No. 4.3e-78;		
Matches 135;	Conservative	0;	Mismatches	1;
			Indels	0;
			Gaps	0;
Qy	1	SSSDKGYKKGGDASFYEP	TPGYLMVNVTGVDKGNELLSPHYVEPIKPGTTLTKKI	60
Db	2	SSSDKGYKKGGDASFYEP	TPGYLMVNVTGVDKGNELLSPHYVEPIKPGTTLTKKI	61
Qy	61	EYYVEWALDATAYKEFRV	ELDPSAKIEVYYDKNKKKEETKSPITEKGFVVPDLSHI	120
Db	62	EYYVEWALDATAYKEFRV	ELDPSAKIEVYYDKNKKKEETKSPITEKGFVVPDLSHI	121
Qy	121	KNPGFNLIITKVIEKK	136	
Db	122	KNPGFNLIITKVIEKK	137	

```

RESULT      6
US-08-852-299-4
; Sequence 4, Application US/08852299
; Patent No. 6010897
; GENERAL INFORMATION:
; APPLICANT: Behnke, Detlef
; APPLICANT: Schlott, Bernhard
; APPLICANT: Albrecht, Sybille
; APPLICANT: G hrs, Karl-Heinz
; APPLICANT: Hartmann, Manfred
; TITLE OF INVENTION: Expression of signal-peptide-free
; staphylokinases
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,299
; FILING DATE: 17-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/256,261
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 137 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-852-299-4

```

Query Match	99.4%	Score 711;	DB 3;	Length 137;
Best Local Similarity	99.3%;	Pred. No. 4.3e-78;		
Matches 135; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

  

Qy	1	SSSFDKGYKGGDASFEPTGPLYMNVNTGVDGSKGELLSPHYVEPIKPGTTLTKEKI	60
Db	2	SSSFDKGYKGGDASFEPTGPLYMNVNTGVDGSKGELLSPHYVEPIKPGTTLTKEKI	61
Qy	61	EYYEWALDATAYKEFRVVELDPSAKTEVYYDNKKKEETKSPITEKGVVPDLSHI	120
Db	62	EYYEWALDATAYKEFRVVELDPSAKTEVYYDNKKKEETKSPITEKGVVPDLSHI	121
Qy	121	KNPGFNLIITKVVIKK	136
Db	122	KNPGFNLIITKVVIKK	137

RESULT 7  
US-08-075-545-1  
: Sequence 1, Application US/08075545  
: Patent No. 5475089  
: GENERAL INFORMATION:  
: APPLICANT: MATSUO, Osamu; SAKAI, Masashi; SHIMURA,  
: APPLICANT: Kisaku; SANSAMA, Hiroshi; WATANABE,  
: APPLICANT: Tsunekazu; MATSUMOTO, Tsuneo; SHISHIDO,  
: APPLICANT: Yoshiyuki; HASHIMOTO, Shusuke; YOKOKURA,  
: APPLICANT: Tetuo; ONOUE, Masaharu; SAKO, Tomoyuki  
: TITLE OF INVENTION: THROMBOLYTIC AGENT  
: NUMBER OF SEQUENCES: 2  
: CORRESPONDENCE ADDRESS:

```

; ADDRESSEE: KECK, MAHIN & CATE
; STREET: P.O. BOX 06110
; CITY: CHICAGO
; STATE: ILLINOIS
; COUNTRY: U.S.A.
; ZIP: 60606-0110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS/PC DOS
; SOFTWARE: DOS Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/075,545
; FILING DATE: 14-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP91/01722
; FILING DATE: 17-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fleit, Martin; Gollin, Michael A.
; REGISTRATION NUMBER: 16,900; 31,957
; REFERENCE/DOCKET NUMBER: 47004-015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 789-3400
; TELEFAX: (202) 789-1158
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 163 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-075-545-1

Query Match      98.68; Score 705; DB 1; Length 163;
Best Local Similarity 98.58; Pred. No. 3e-77;
Matches 134; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSSFDKGYKKGGDASVFPTGPLYMNVNVTGVDKGNELLSPHYVEPIKPGTTLTKEKI 60
DB 28 SSSFDKGYKKGGDASVFPTGPLYMNVNVTGVDKGNELLSPHYVEPIKPGTTLTKEKI 87
QY 61 EYVVEWALDATAYKEFRVVELDPSAKIEVTYYDKNKKKEETKSPITEKGFVVPDLSEHI 120
DB 88 EYVVEWALDATAYKEFRVVELDPSAKIEVTYYDKNKKKEETKSPITEKGFVVPDLSEHI 147
QY 121 KNPGFNLTITKVIEKK 136
DB 148 KNPGFNLTITKVIEKK 163

RESULT 8
US-08-256-261-2
; Sequence 2, Application US/08256261
; Patent No. 5801037
; GENERAL INFORMATION:
; APPLICANT: Behnke, Detlef
; APPLICANT: Schlott, Bernhard
; APPLICANT: Albrecht, Sybille
; APPLICANT: G hrs, Karl-Heinz
; APPLICANT: Hartmann, Manfred
; TITLE OF INVENTION: Expression of signal-peptide-free
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,299
; FILING DATE: 17-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/256,261
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 137 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; ADDRESSEE: KECK, MAHIN & CATE
; STREET: P.O. BOX 06110
; CITY: CHICAGO
; STATE: ILLINOIS
; COUNTRY: U.S.A.
; ZIP: 60606-0110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS/PC DOS
; SOFTWARE: DOS Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/075,545
; FILING DATE: 14-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP91/01722
; FILING DATE: 17-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fleit, Martin; Gollin, Michael A.
; REGISTRATION NUMBER: 16,900; 31,957
; REFERENCE/DOCKET NUMBER: 47004-015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 789-3400
; TELEFAX: (202) 789-1158
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 163 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-075-545-1

Query Match      97.28; Score 695; DB 1; Length 137;
Best Local Similarity 97.88; Pred. No. 3.7e-76;
Matches 133; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSFDKGYKKGGDASVFPTGPLYMNVNVTGVDKGNELLSPHYVEPIKPGTTLTKEKI 60
DB 2 SSSFDKGYKKGGDASVFPTGPLYMNVNVTGVDKGNELLSPHYVEPIKPGTTLTKEKI 61
QY 61 EYVVEWALDATAYKEFRVVELDPSAKIEVTYYDKNKKKEETKSPITEKGFVVPDLSEHI 120
DB 62 EYVVEWALDATAYKEFRVVELDPSAKIEVTYYDKNKKKEETKSPITEKGFVVPDLSEHI 121
QY 121 KNPGFNLTITKVIEKK 136
DB 122 KNPGFNLTITKVIEKK 137

RESULT 9
US-08-852-299-2
; Sequence 2, Application US/08852299
; Patent No. 6010897
; GENERAL INFORMATION:
; APPLICANT: Behnke, Detlef
; APPLICANT: Schlott, Bernhard
; APPLICANT: Albrecht, Sybille
; APPLICANT: G hrs, Karl-Heinz
; APPLICANT: Hartmann, Manfred
; TITLE OF INVENTION: Expression of signal-peptide-free
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,299
; FILING DATE: 17-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/256,261
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 137 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-852-299-2
```



```
Query Match          97.2%; Score 695; DB 3; Length 137;
Best Local Similarity 97.8%; Pred. No. 3.7e-76;
Matches 133; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSFDKGYKKGGDASYFEPTGYPLMVNVTGVDKGNELLSPHYVEFPKPGTTLTKKI 60
   |||||
Db 2 SSSFDKGYKKGGDASYFEPTGYPLMVNVTGVDKGNELLSPHYVEFPKPGTTLTKKI 61
   |||||

QY 61 EYVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSFPIPEKGFVVPDLSEHI 120
   |||||
Db 62 EYVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSFPIPEKGFVVPDLSEHI 121
   |||||

QY 121 KNPGFNLTWKVIEKK 136
   |||||
Db 122 KNPGFNLTWKVIEKK 137
   |||||

RESULT 10
US-08-256-261-17
; Sequence 17, Application US/08256261
; Patent No. 5801037
; GENERAL INFORMATION:
; APPLICANT: Behnke, Detlef
; APPLICANT: Schlott, Bernhard
; APPLICANT: Albrecht, Sybille
; APPLICANT: G hrs, Karl-Heinz
; APPLICANT: Hartmann, Manfred
; TITLE OF INVENTION: Expression of signal-peptide-free
; TITLE OF INVENTION: staphylokinases
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,261
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 163 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-256-261-17

Query Match          97.2%; Score 695; DB 1; Length 163;
Best Local Similarity 97.8%; Pred. No. 4.7e-76;
Matches 133; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSFDKGYKKGGDASYFEPTGYPLMVNVTGVDKGNELLSPHYVEFPKPGTTLTKKI 60
   |||||
Db 28 SSSFDKGYKKGGDASYFEPTGYPLMVNVTGVDKGNELLSPHYVEFPKPGTTLTKKI 87
   |||||

QY 61 EYVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSFPIPEKGFVVPDLSEHI 120
   |||||
Db 88 EYVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSFPIPEKGFVVPDLSEHI 147
   |||||

QY 121 KNPGFNLTWKVIEKK 136
   |||||
Db 148 KNPGFNLTWKVIEKK 163
   |||||
```

```
RESULT 11
US-08-852-299-17
; Sequence 17, Application US/08852299
; Patent No. 6010897
; GENERAL INFORMATION:
; APPLICANT: Behnke, Detlef
; APPLICANT: Schlott, Bernhard
; APPLICANT: Albrecht, Sybille
; APPLICANT: G hrs, Karl-Heinz
; APPLICANT: Hartmann, Manfred
; TITLE OF INVENTION: Expression of signal-peptide-free
; TITLE OF INVENTION: staphylokinases
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,299
; FILING DATE: 17-MAY-1997
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 163 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-852-299-17

Query Match          97.2%; Score 695; DB 3; Length 163;
Best Local Similarity 97.8%; Pred. No. 4.7e-76;
Matches 133; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSFDKGYKKGGDASYFEPTGYPLMVNVTGVDKGNELLSPHYVEFPKPGTTLTKKI 60
   |||||
Db 28 SSSFDKGYKKGGDASYFEPTGYPLMVNVTGVDKGNELLSPHYVEFPKPGTTLTKKI 87
   |||||

QY 61 EYVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSFPIPEKGFVVPDLSEHI 120
   |||||
Db 88 EYVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSFPIPEKGFVVPDLSEHI 147
   |||||

QY 121 KNPGFNLTWKVIEKK 136
   |||||
Db 148 KNPGFNLTWKVIEKK 163
   |||||

RESULT 12
US-08-256-261-14
; Sequence 14, Application US/08256261
; Patent No. 5801037
; GENERAL INFORMATION:
; APPLICANT: Behnke, Detlef
; APPLICANT: Schlott, Bernhard
; APPLICANT: Albrecht, Sybille
; APPLICANT: G hrs, Karl-Heinz
; APPLICANT: Hartmann, Manfred
; TITLE OF INVENTION: Expression of signal-peptide-free
; TITLE OF INVENTION: staphylokinases
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
```

```
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,261
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 137 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-256-261-14

Query Match          96.8%; Score 692; DB 1; Length 137;
Best Local Similarity 97.1%; Pred. No. 8.4e-76;
Matches 132; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSFDKGYKKGDDASYFEPTGPLYMNVNVTGVDKGNELLSPHYVEFPKPGTTLTKKI 60
Db 2 SSSFDKGYKKGDDASYFEPTGPLYMNVNVTGVDKGNELLSPHYVEFPKPGTTLTKKI 61
QY 61 EYVVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSPFITEKGFVVPDLSEHI 120
Db 62 EYVVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSPFITEKGFVVPDLSEHI 121
QY 121 KNPGFNLTIKVVEKK 136
Db 122 KNPGFNLTIKVVEKK 137

RESULT 13
US-08-256-299-14
; Sequence 14, Application US/08852299
; Patent No. 6010897
; GENERAL INFORMATION:
; APPLICANT: Behnke, Detlef
; APPLICANT: Schlott, Bernhard
; APPLICANT: Albrecht, Sybille
; APPLICANT: G hrs, Karl-Heinz
; APPLICANT: Hartmann, Manfred
; TITLE OF INVENTION: Expression of signal-peptide-free
; TITLE OF INVENTION: staphylokinases
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,299
; FILING DATE: 17-MAY-1997
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 137 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-256-261-14

Query Match          96.8%; Score 692; DB 1; Length 137;
Best Local Similarity 97.1%; Pred. No. 8.4e-76;
Matches 132; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSFDKGYKKGDDASYFEPTGPLYMNVNVTGVDKGNELLSPHYVEFPKPGTTLTKKI 60
Db 2 SSSFDKGYKKGDDASYFEPTGPLYMNVNVTGVDKGNELLSPHYVEFPKPGTTLTKKI 61
QY 61 EYVVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSPFITEKGFVVPDLSEHI 120
Db 62 EYVVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSPFITEKGFVVPDLSEHI 121
QY 121 KNPGFNLTIKVVEKK 136
Db 122 KNPGFNLTIKVVEKK 137

RESULT 13
US-08-256-299-14
; Sequence 14, Application US/08852299
; Patent No. 6010897
; GENERAL INFORMATION:
; APPLICANT: Behnke, Detlef
; APPLICANT: Schlott, Bernhard
; APPLICANT: Albrecht, Sybille
; APPLICANT: G hrs, Karl-Heinz
; APPLICANT: Hartmann, Manfred
; TITLE OF INVENTION: Expression of signal-peptide-free
; TITLE OF INVENTION: staphylokinases
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,261
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 137 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-256-261-14
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 137 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-852-299-14

Query Match          96.4%; Score 689; DB 1; Length 137;
Best Local Similarity 97.1%; Pred. No. 1.9e-75;
Matches 132; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSSFDKGYKKGDDASYFEPTGPLYMNVNVTGVDKGNELLSPHYVEFPKPGTTLTKKI 60
Db 2 SSSFDKGYKKGDDASYFEPTGPLYMNVNVTGVDKGNELLSPHYVEFPKPGTTLTKKI 61
QY 61 EYVVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSPFITEKGFVVPDLSEHI 120
Db 62 EYVVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSPFITEKGFVVPDLSEHI 121
QY 121 KNPGFNLTIKVVEKK 136
Db 122 KNPGFNLTIKVVEKK 137

RESULT 14
US-08-256-261-12
; Sequence 12, Application US/08256261
; Patent No. 5801037
; GENERAL INFORMATION:
; APPLICANT: Behnke, Detlef
; APPLICANT: Schlott, Bernhard
; APPLICANT: Albrecht, Sybille
; APPLICANT: G hrs, Karl-Heinz
; APPLICANT: Hartmann, Manfred
; TITLE OF INVENTION: Expression of signal-peptide-free
; TITLE OF INVENTION: staphylokinases
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,261
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 137 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-256-261-12

Query Match          96.4%; Score 689; DB 1; Length 137;
Best Local Similarity 97.1%; Pred. No. 1.9e-75;
Matches 132; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSSFDKGYKKGDDASYFEPTGPLYMNVNVTGVDKGNELLSPHYVEFPKPGTTLTKKI 60
Db 2 SSSFDKGYKKGDDASYFEPTGPLYMNVNVTGVDKGNELLSPHYVEFPKPGTTLTKKI 61
QY 61 EYVVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSPFITEKGFVVPDLSEHI 120
Db 62 EYVVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSPFITEKGFVVPDLSEHI 121
QY 121 KNPGFNLTIKVVEKK 136
Db 122 KNPGFNLTIKVVEKK 137

RESULT 14
US-08-256-261-12
; Sequence 12, Application US/08256261
; Patent No. 5801037
; GENERAL INFORMATION:
; APPLICANT: Behnke, Detlef
; APPLICANT: Schlott, Bernhard
; APPLICANT: Albrecht, Sybille
; APPLICANT: G hrs, Karl-Heinz
; APPLICANT: Hartmann, Manfred
; TITLE OF INVENTION: Expression of signal-peptide-free
; TITLE OF INVENTION: staphylokinases
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,261
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 137 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-256-261-12

Query Match          96.4%; Score 689; DB 1; Length 137;
Best Local Similarity 97.1%; Pred. No. 1.9e-75;
Matches 132; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSSFDKGYKKGDDASYFEPTGPLYMNVNVTGVDKGNELLSPHYVEFPKPGTTLTKKI 60
Db 2 SSSFDKGYKKGDDASYFEPTGPLYMNVNVTGVDKGNELLSPHYVEFPKPGTTLTKKI 61
QY 61 EYVVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSPFITEKGFVVPDLSEHI 120
Db 62 EYVVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSPFITEKGFVVPDLSEHI 121
QY 121 KNPGFNLTIKVVEKK 136
Db 122 KNPGFNLTIKVVEKK 137

RESULT 14
US-08-256-261-12
; Sequence 12, Application US/08256261
; Patent No. 5801037
; GENERAL INFORMATION:
; APPLICANT: Behnke, Detlef
; APPLICANT: Schlott, Bernhard
; APPLICANT: Albrecht, Sybille
; APPLICANT: G hrs, Karl-Heinz
; APPLICANT: Hartmann, Manfred
; TITLE OF INVENTION: Expression of signal-peptide-free
; TITLE OF INVENTION: staphylokinases
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,261
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 137 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-256-261-12
```

Search completed: April 22, 2002, 10:41:49  
Job time: 90 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 10:38:39 ; Search time 14.37 Seconds  
(without alignments)  
720.928 Million cell updates/sec

Title: US-09-601-490-1  
Perfect score: 715  
Sequence: 1 SSSFDKGYKGGDASYPF.....SEHIKNPGFNLTITKVIIEKK 136

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	711	99.4	163	1 PRSAK	staphylokinase - p
2	695	97.2	163	2 S02330	staphylokinase - p
3	87	12.2	348	2 B69790	hypothetical prote
4	84.5	11.8	172	1 WMVZTH	BamHI-ORF13 protei
5	83	11.6	611	2 D72275	conserved hypothet
6	80.5	11.3	158	2 B71192	hypothetical prote
7	80.5	11.3	3097	2 T00021	DN-cadherin - frui
8	78.5	11.0	324	2 E64487	hypothetical prote
9	78	10.9	350	2 C56118	vetispiradiene syn
10	77.5	10.8	441	1 ITBPT4	DNA topoisomerase
11	77	10.8	559	2 T08174	sesquiterpene cycl
12	77	10.8	608	2 I53269	prolactin receptor
13	76.5	10.7	289	2 C69349	conserved hypothet
14	76.5	10.7	506	1 PIWUB4	L1 protein - bovin
15	76	10.6	252	2 E86822	hypothetical prote
16	75.5	10.6	327	2 C35411	alkanal monooxygen
17	75	10.5	576	2 G72277	NH(3)-dependent NA
18	74.5	10.4	2512	1 XYCHFA	fatty-acid synthas
19	74	10.3	233	2 T24714	hypothetical prote
20	74	10.3	233	2 D69407	medium-chain acyl-
21	74	10.3	256	2 T24713	hypothetical prote
22	74	10.3	298	2 H64402	molybdopterin bios
23	74	10.3	462	2 T01732	UTP-glucose glucos
24	74	10.3	480	2 B85014	probable flavonol
25	74	10.3	552	2 D82878	phosphomannomutase
26	74	10.3	610	2 A34631	lactogen receptor
27	74	10.3	630	2 A36116	prolactin receptor
28	73.5	10.3	430	2 T04668	phosphoserine tran
29	73.5	10.3	431	2 E96671	hypothetical prote

30	73.5	10.3	764	2 T39878	hypothetical prote
31	73.5	10.3	1536	2 E72310	hypothetical prote
32	73	10.2	237	2 C84976	tRNA (guanine-N1-)
33	73	10.2	263	2 C63044	hypothetical prote
34	73	10.2	470	2 B96637	hypothetical prote
35	73	10.2	2505	1 XYRTFA	fatty-acid synthas
36	73	10.2	2509	2 G01880	fatty-acid synthas
37	73	10.2	4969	2 A37113	ryanodine receptor
38	72.5	10.1	374	2 C85791	hypothetical prote
39	72.5	10.1	374	2 B64941	probable choline m
40	72.5	10.1	430	2 B84946	peptidylprolyl iso
41	72.5	10.1	584	2 C70425	transporter (extra
42	72.5	10.1	935	2 T50774	probable translati
43	72.5	10.1	1043	2 F86504	isoleucyl-tRNA syn
44	72.5	10.1	1043	2 E72118	isoleucine--tRNA
45	72	10.1	417	2 A75563	2-oxoglutarate deh

ALIGNMENTS

RESULT 1

PRSAK

staphylokinase - phage S phi-C

C:Species: phage S phi-C

A:Note: host Staphylococcus aureus

C:Date: 28-Aug-1985 #sequence\_revision 28-Aug-1985 #text\_change 28-May-1999

C:Accession: A00995

R:Sako, T.; Tsuchida, N.

Nucleic Acids Res. 11, 7679-7693, 1983

A:Title: Nucleotide sequence of the staphylokinase gene from Staphylococcus aureus.

A:Reference number: A00995; MUID:84069795

A:Accession: A00995

A:Molecule type: DNA

A:Residues: 1-163 <SAK>

A:Cross-references: GB:X00127; NID:g47425; PIDN:CAA24957.1; PID:g758303

C:Comment: Although it has no intrinsic proteolytic activity, this secreted protein i

cture of plasminogen is thought to expose the plasmin active site.

C:Comment: The designation of staphylokinase as synonomous with Staphylococcal aureus

C:Genetics:

A:Gene: sak

C:Superfamily: phage S phi-C staphylokinase

C:Keywords: plasminogen activator

Query Match 99.4%; Score 711; DB 1; Length 163;  
Best Local Similarity 99.3%; Pred. No. 7.8e-57;  
Matches 135; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSSFDKGYKGGDASYPFPTGPLYMNVGTGDSKGNELLSPHYVEFPKPGTTLTKEKI 60

Db 28 SSSFDKGYKGGDASYPFPTGPLYMNVGTGDSKGNELLSPHYVEFPKPGTTLTKEKI 87

QY 61 EYVVEWALDATAYKEFRVWELDSPAKIEVTYYDKNKKKEETKSPITEKGVVPDLSEHI 120

Db 88 EYVVEWALDATAYKEFRVWELDSPAKIEVTYYDKNKKKEETKSPITEKGVVPDLSEHI 147

QY 121 KNPGFNLTITKVIIEKK 136

Db 148 KNPGFNLTITKVIIEKK 163

RESULT 2

S02330

staphylokinase - phage P42D

C:Species: phage P42D

A:Note: host Staphylococcus aureus

C:Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 26-Aug-1999

C:Accession: S02330; S45654

R:Behnke, D.; Gerlach, D.

Mol. Gen. Genet. 210, 528-534, 1987

A:Title: Cloning and expression in Escherichia coli, Bacillus subtilis, and Streptoco

A:Reference number: S02330; MUID:88121731

A:Accession: S02330  
A:Molecule type: DNA  
A:Residues: 1-163 <BEH>  
A:Cross-references: EMBL:X06603; NID:g46676; PIDN:CAA29822.1; PID:g758275  
R:Gase, A.; Birch-Hirschfeld, E.; Guehrs, K.H.; Hartmann, M.; Vetterman, S.; Damaschun, Eur. J. Biochem. 223, 303-308, 1994  
A:Title: The thermostability of natural variants of bacterial plasminogen-activator stad  
A:Reference number: S45654; MUID:94307274  
A:Accession: S45654  
A:Molecule type: protein  
A:Residues: 28-163 <AS>  
C:Genetics:  
A:Gene: sak  
C:Superfamily: phage S phi-C staphylokinase

Query Match 97.2%; Score 695; DB 2; Length 163;  
Best Local Similarity 97.8%; Pred. No. 2.1e-55;  
Matches 133; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSFDKGYKKGDDASVFETPGPYLMVNVTVGVDKGNELLSPHYVFPPIKPGTTLTKEKI 60  
|||||  
Db 28 SSSFDKGYKKGDDASVFETPGPYLMVNVTVGVDKGNELLSPHYVFPPIKPGTTLTKEKI 60  
|||||

QY 61 EYVWEALDATAKERVVVELDPSAKIEVYYDKNKKKEETKSPFITERKGFVVPDLSEHI 120  
|||||  
Db 88 EYVWEALDATAKERVVVELDPSAKIEVYYDKNKKKEETKSPFITERKGFVVPDLSEHI 147  
|||||

QY 121 KNPGFNLITKVIEKK 136  
|||||  
Db 148 KNPGFNLITKVIEKK 163  
|||||

RESULT 3  
B69790  
Hypothetical protein ydJN - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 15-Oct-1999  
C:Accession: B69790  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertei, C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choc, A.; Ehrlich, S.D.; Emmeron, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.; Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueh, Y.M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, R.; Rieger, M.; Rivoita, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot, T.; Winters, P.; Wipat, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, A.; Winters, P.; Wipat, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, A.; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A69580; MUID:98044033  
A:Accession: B69790  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-348 <UN>  
A:Cross-references: GB:299107; GB:AL009126; NID:g2632866; PIDN:CAB12445.1; PID:ell182605;  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: ydJN

Query Match 12.2%; Score 87; DB 2; Length 348;  
Best Local Similarity 26.5%; Pred. No. 2.1;  
Matches 31; Conservative 18; Mismatches 34; Indels 34; Gaps 5;

QY 5 DKGYKKGDDASVFETPGPYLMVNVTVGVDKGNELLSPHYVFPPIKPGTTLTKEIYYV 64  
|||||  
Db 27 DKG-HEKAADVNTSGDELLVSDTD-----LLTKY-----ENDKVIHEEKLTSYP 75  
|||||

QY 65 EWALDATAKERVVVELDPSAKIEVYYDKNKKKEETKSPFITERKGFVVPDLSEHIK 121  
|||||  
Db 76 AFALD-----KQQVLYTGNNEQNEMLFKL-----DLKSHKK 109  
|||||

RESULT 4  
WMVZTH  
BamHI-ORF13 protein - fowlpox virus (isolate HP-438[Munich]) (fragment)  
C:Species: fowlpox virus  
C:Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 16-Jun-2000  
C:Accession: D30087  
R:Tomley, F.; Binns, M.; Campbell, J.; Boursnell, M.  
J. Gen. Virol. 69, 1025-1040, 1988  
A:Title: Sequence analysis of an 11.2 kilobase, near-terminal, BamHI fragment of fowl  
A:Reference number: JT0442; MUID:88229622  
A:Accession: D30087  
A:Molecule type: DNA  
A:Residues: 1-172 <TOM>  
A:Cross-references: GB:D00295; NID:g221380; PIDN:BAA00209.1; PID:g221400  
C:Superfamily: fowlpox virus BamHI-ORF13 protein  
C:Keywords: early protein

Query Match 11.8%; Score 84.5; DB 1; Length 172;  
Best Local Similarity 25.2%; Pred. No. 1.5;  
Matches 31; Conservative 19; Mismatches 36; Indels 37; Gaps 5;

QY 43 HYVEFPIKPGTTLTKEKI-EYVWEALDATA-----YKEFRV-----VEL 81  
|||||  
Db 38 HYCLKPIHIAANRTESKIVKLLIEGADINSEGGANGKYPIHYAMKVYDPFLKGIKVL 97  
|||||

QY 82 DPSAKIEVYYDKNKKKEETKSPFITERKGFVVPDLSEHIKNPGLN-----ITKVV 132  
|||||  
Db 98 DHGA-----DINKQSVLTNTSPLYETRFITDLDLVIISRGANINKEEWEVITRII 150  
|||||

QY 133 IEK 135  
|||  
Db 151 LER 153  
|||

RESULT 5  
D72275  
conserved hypothetical protein - Thermotoga maritima (strain MS88)  
C:Species: Thermotoga maritima  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: D72275  
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hic, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, C.M.  
Nature 399, 323-329, 1999  
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome  
A:Reference number: A72200; MUID:99287316  
A:Accession: D72275  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-611 <ARN>  
A:Cross-references: GB:AE001781; GB:AE000512; NID:g4981810; PIDN:AAD36331.1; PID:g498  
A:Experimental source: strain MS88  
C:Genetics:  
A:Gene: TM1257

Query Match 11.6%; Score 83; DB 2; Length 611;  
Best Local Similarity 21.5%; Pred. No. 9.3;  
Matches 28; Conservative 27; Mismatches 45; Indels 30; Gaps 5;

QY 8 KYKKGDDASYFFPTG-----PYLMVNVTVGVDKGN-----ELLSPHYVFPPIKP 51  
|||||  
Db 163 RYILGDGAHINVSQVAAKTSYTFVLKSNITSSKNGDLMREARAYIIFNVAK 222  
|||||

QY 52 GTTLTKEIYYVEALDATAKKE-FRVVELDPSAKIEVYYDKNKKKEETKSPFITERK 110  
|||||  
Db 223 ESSLFLDRIS--KEWYSEREKWDENRYVLGIEPKPFENVAFY-----AFSREKG 269  
|||||



Query Match 10.9%; Score 78; DB 2; Length 350;  
Best Local Similarity 32.8%; Pred. No. 13;  
Matches 22; Conservative 14; Mismatches 23; Indels 8; Gaps 4;  
QY 59 KIEYVEWALDAYAYKEFRVVELDPSAKIEVTVYDKNKKKEETKSPITEKGFV---VPD 115  
Db 138 KISKYV--LLD--LYKDYE--TELSKDGSRSEVHYAKERKEIVRNFFVEAKWFIEGYMPP 192  
QY 116 LSEHIKN 122  
Db 193 VSEYLN 199  
RESULT 10  
ITPBT4  
Nucleic Acids Res. 14, 7379-7390, 1986  
A:Title: The 52-protein subunit of T4 DNA topoisomerase is homologous to the gyra-protein  
A:Reference number: A24705; MUID:87016377  
A:Accession: B24705  
A:Molecule type: DNA  
A:Residues: 1-441 <HUA>  
R:Chapman, D.; Morad, I.; Kaufmann, G.; Gait, M.J.; Jorissen, L.; Snyder, L.  
J. Mol. Biol. 199, 373-377, 1988  
A:Title: Nucleotide and deduced amino acid sequence of stp: the bacteriophage T4 anticom  
A:Reference number: S01711; MUID:88172481  
A:Accession: S01872  
A>Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-46 <CHA>  
C:Genetics:  
A:Gene: 52  
A:Map position: 162-164 kb  
A:Complex: this is the middle-sized of the three subunits of T4 DNA topoisomerase  
C:Superfamily: phase T4 DNA topoisomerase (ATP-hydrolyzing) medium chain; phase T4 DNA  
C:Keywords: DNA binding; DNA replication; isomerase  
F:1-225/Domain: phase T4 DNA topoisomerase (ATP-hydrolyzing) medium chain homology <TOP>  
Query Match 10.8%; Score 77.5; DB 1; Length 441;  
Best Local Similarity 25.5%; Pred. No. 20;  
Matches 36; Conservative 21; Mismatches 43; Indels 41; Gaps 8;  
QY 16 SYFEPTGPLYMVN-VTGVDKSGNELLSPHYVEFFIK-----PGTTLTKETKIEYVVEWAL 68  
Db 147 AFYLPITPLVLLVSGIATGATVILPHSVSSVKKAVLQALOGKKVTKPVE----- 199  
QY 69 DATAYKEFR--VVELDPSAKIEVTVYDKNKKKEETKSF-PITEKGF 111  
Db 200 -----EPEFRGEVVEIDGVEINGTYKFTSRQTHITEIPYKYDRETYVYSKILDPLENKGF 255  
QY 112 VVPD--LSEHIKNPGFNLT 130  
Db 256 TWDDACGEH----GFGFKVK 272  
RESULT 11  
T08174  
sesquiterpene cyclase (EC 2.5.1.-) - pepper  
C:Species: Capsicum annuum (pepper)  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: T08174  
R:Back, K.; Shin, D.H.; He, S.  
Plant Cell Physiol. 39, 899-904, 1998  
A:Title: Cloning and bacterial expression of sesquiterpene cyclase, a key branch point e  
A:Reference number: Z16395; MUID:99033462

A:Accession: T08174  
A>Status: Preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-559 <BAC>  
A:Cross-references: EMBL:AF061285; NID:g3108342; PIDN:AAC61260.1; PID:g3108343  
C:Function:  
A:Pathway: the synthesis of phytoalexin capsidiol  
A>Note: unduced by UV  
C:Superfamily: vetispiradiene synthase 1  
C:Keywords: isoprenoid biosynthesis; transferase  
Query Match 10.8%; Score 77; DB 2; Length 559;  
Best Local Similarity 34.0%; Pred. No. 29;  
Matches 18; Conservative 11; Mismatches 20; Indels 4; Gaps 2;  
QY 73 YKEFRVVELDPSAKIEVTVYDKNKKKEETKSPITEKGFV---VPDLSEHIKN 122  
Db 357 YKDYE--KEMSRDGRSHVVYAKERLKLKELVKSYNIEAKWFIEGHMPPASEYLRN 408  
RESULT 12  
I53269  
prolactin receptor, long form - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Aug-1996 #Sequence revision 02-Aug-1996 #text\_change 28-Jul-2000  
C:Accession: I53269; JT0671; S34356  
R:Clarke, D.L.; Linzer, D.I.H.  
Endocrinology 133, 224-232, 1993  
A:Title: Changes in prolactin receptor expression during pregnancy in the mouse ovary  
A:Reference number: I53269; MUID:93307149  
A:Accession: I53269  
A>Status: Preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-608 <RES>  
A:Cross-references: GB:L14811; NID:g293769; PIDN:AAA02686.1; PID:g293770  
R:Moore, R.C.; Oka, T.  
Gene 134, 263-265, 1993  
A:Title: Cloning and sequencing of the cDNA encoding the murine mammary gland long-f  
A:Reference number: JT0671; MUID:94085788  
A:Accession: JT0671  
A:Molecule type: mRNA  
A:Residues: 1-608 <MOU>  
A:Cross-references: GB:L13593; NID:g347398; PIDN:AAC37641.1; PID:g347842  
R:Edery, M.; Pezet, A.; Nandi, S.; Kelly, P.A.  
submitted to the EMBL Data Library, June 1993  
A:Description: Isolation and nucleotide sequence of a mouse cDNA prolactin receptor.  
A:Reference number: S34356  
A:Accession: S34356  
A:Molecule type: mRNA  
A:Residues: 1-557 'F', 559-608 <EDE>  
A:Cross-references: EMBL:X73372; NID:g312696; PIDN:CAA51789.1; PID:g312697  
C:Comment: Prolactin receptor have long form and short form which are resulted from a  
C:Comment: This long form receptor is capable of transducing a signal to milk protein  
C:Superfamily: cytokine receptor homology  
C:Keywords: receptor; transmembrane protein  
F:31-216/Domain: cytokine receptor homology <CRS>  
F:230-253/Domain: transmembrane #status predicted <TMM>  
Query Match 10.8%; Score 77; DB 2; Length 608;  
Best Local Similarity 23.9%; Pred. No. 32;  
Matches 34; Conservative 17; Mismatches 43; Indels 48; Gaps 7;  
QY 4 FDGKYKKGDDASY-----FEPTGYPY--LMNVNTGVDSKGNELLSLP-HYVEFP---I 49  
Db 274 FDTLLEKKGSEELLSALGCDQDFPTSDCDLLVFEFLVEDDNEDERLMPSHSKEYPGGV 333  
QY 50 KPGTTLTKETKIEYVVEWALDAYAYKEFRVVELDPSAKIEVTVYDKNKKKEETKSF 105  
Db 334 KP-----THLDPDSGSHGSDSHSLLSLCKEPEQAYP 366  
QY 106 ITEKGFVVVPLDSEHIKNPGFN 127







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 10:41:34 ; Search time 11.66 Seconds  
(without alignments)  
427.651 Million cell updates/sec

Title: US-09-601-490-1  
Perfect score: 715  
Sequence: 1 SSSFDKGYKGGDASIFEP.....SEHIKNPGFNLTIVKVIKK 136

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	711	99.4	163	SAK_STAAU	P00802 staphylococ
2	695	97.2	163	SAK_BPP42	P15240 bacterioph
3	84.5	11.8	172	YB13_FOWPM	P14368 fowlpox vir
4	80.5	11.3	3097	CADN_DROME	O15943 drosophila
5	78.5	11.0	324	YF02_METJA	Q58897 methanococ
6	78.5	11.0	719	IF39_TOBAC	P56821 nicotiana t
7	77.5	10.8	442	TOP5_BPT4	P07065 bacterioph
8	77	10.8	608	PLRL_MOUSE	Q08501 mus musculu
9	76.5	10.7	506	VLI_BPV4	P08341 bovine papi
10	75.5	10.6	327	LXBI_PHOLU	P19840 photorhabdu
11	75	10.5	576	NAE2_THEMA	Q9x0y0 thermotoga
12	74.5	10.4	2511	FAS_CHICK	P12276 gallus gall
13	74	10.3	298	MORA_METJA	Q58234 methanococ
14	74	10.3	610	PLRL_RAT	P05710 rattus norv
15	73.5	10.3	430	SERC_ARATH	Q96255 arabidopsis
16	73	10.2	231	CTRA_CAUCR	Q45994 caulobacter
17	73	10.2	237	TRMD_BUCAI	P57476 buchnera ap
18	73	10.2	263	YPUB_ECOLI	P76612 escherichia
19	73	10.2	2505	FAS_RAT	P12785 rattus norv
20	73	10.2	4969	RYNC_RABIT	P30957 cryptotaguis
21	72.5	10.1	374	YEAU_ECOLI	P76253 escherichia
22	72.5	10.1	430	SURA_BUCAI	P57240 buchnera ap
23	72.5	10.1	935	IF38_MEDTR	Q9xhml medicago tr
24	72.5	10.1	1043	SYI_CHLPN	Q92972 chlamydia p
25	72	10.1	429	HISX_METJA	Q58851 methanococ
26	72	10.1	695	APP2_MOUSE	Q06335 mus musculu
27	72	10.1	978	KFMS_RAT	Q00495 rattus norv
28	72	10.1	1102	YE20_METJA	Q58815 methanococ
29	72	10.1	2663	CENE_HUMAN	Q02224 homo sapien
30	71.5	10.0	553	YF61_METJA	Q58956 methanococ
31	71.5	10.0	560	POTA_MYCPN	P75059 mycoplasma
32	71.5	10.0	667	Y366_MYCGE	P47606 mycoplasma
33	71.5	10.0	686	FREL_YEAST	P32791 saccharomyc

ALIGNMENTS

RESULT 1

ID	SAK_STAAU	STANDARD;	PRT;	163 AA.
AC	P00802;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	STAPHYLOKINASE PRECURSOR (NEUTRAL PROTEINASE) (PROTEASE III).			
GN	SAK.			
OS	Staphylococcus aureus.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;			
OC	Bacillus/Staphylococcus group; Staphylococcus.			
OX	NCBI_TaxID=1280;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=84069795; PubMed=6359061;			
RA	Sako T., Tsuchida N.;			
RT	"Nucleotide sequence of the staphylokinase gene from Staphylococcus aureus.";			
RL	Nucleic Acids Res. 11:7679-7693(1983).			
RN	[2]			
RP	X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 43-163.			
RX	MEDLINE=97290447; PubMed=9145104;			
RA	Rabljns A., de Bondt H.L., de Ranter C.;			
RT	"Three-dimensional structure of staphylokinase, a plasminogen activator with therapeutic potential.";			
RL	Nat. Struct. Biol. 4:357-360(1997).			
RN	[3]			
RP	STRUCTURE BY NMR OF 28-163.			
RX	MEDLINE=98367505; PubMed=9692953;			
RA	Ohlenschlaeger O., Ramachandran R., Guehrs K.H., Schlott B.,			
RT	"Nuclear magnetic resonance solution structure of the plasminogen-activator protein staphylokinase.";			
RL	Biochemistry 37:10635-10642(1998).			
CC	- - FUNCTION: THIS EXTRACELLULAR PROTEIN IS ONE OF THE PLASMINOGEN			
CC	ACTIVATORS THAT CONVERTS PLASMINOGEN INTO PLASMIN; IT REQUIRES			
CC	CALCIUM ION FOR STABILIZATION.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; X00127; CAA24957.1; -.			
DR	PIR; A00995; PRSAK.			
DR	PDB; 2SAK; 25-FEB-98.			
DR	PDB; 1SSN; 02-DEC-98.			
KW	Hydrolase; Calcium; Plasminogen activation; Signal; 3D-structure.			
FT	SIGNAL 1 27 POTENTIAL.			
FT	CHAIN 28 163 STAPHYLOKINASE.			
SQ	SEQUENCE 163 AA; 18490 MW; E56D9FF50AEDE141 CRC64;			

P14788 synchococc  
Q27367 drosophila  
P26404 salmonella  
P12048 b bifunctio  
Q54873 streptococc  
P09581 mus musculu  
P41885 caenorhabdi  
P39077 saccharomyc  
Q9hhc4 pyrococcus  
Q04399 saccharomyc  
P75432 mycoplasma  
Q57703 methanococc

```
Query Match          99.4%; Score 711; DB 1; Length 163;
Best Local Similarity 99.3%; Pred. No. 8.e-59;
Matches 135; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSSFDKGYKKGDDASFEPTGYLMVNTGVDSKGNELLSPHYVEFPKPGTTLTKKEI 60
   |||||
Db 28 SSSFDKGYKKGDDASFEPTGYLMVNTGVDSKGNELLSPHYVEFPKPGTTLTKKEI 87
   |||||

QY 61 EYVVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSFPITEKGFVVPDLSEHI 120
   |||||
Db 88 EYVVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSFPITEKGFVVPDLSEHI 147
   |||||

QY 121 KNPGENLITKVIEKK 136
   |||||
Db 148 KNPGENLITKVIEKK 163
   |||||

RESULT 2
SAK_BBP42
ID SAK_BBP42 STANDARD; PRT; 163 AA.
AC P15340;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE STAPHYLOKINASE PRECURSOR (NEUTRAL PROTEINASE) (PROTEASE III).
GN SAK.
OS Bacteriophage P42D.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;
OC Lambda phage group.
OX NCBI_TaxID=10715;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88121731; PubMed=3123893;
RA Behnke D., Gerlach D.;
RT "Cloning and expression in Escherichia coli, Bacillus subtilis, and
RT Streptococcus sanguis of a gene for staphylokinase -- a bacterial
RT plasminogen activator.";
RL Mol. Gen. Genet. 210:528-534(1987).
CC -!- FUNCTION: THIS EXTRACELLULAR PROTEIN IS ONE OF THE PLASMINOGEN
CC ACTIVATORS THAT CONVERTS PLASMINOGEN INTO PLASMIN; IT REQUIRES
CC CALCIUM ION FOR STABILIZATION.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X06603; CAA29822.1; -
DR EMBL; M57455; AAA98206.1; -
DR EMBL; A17537; CAA01341.1; -
DR PIR; S02330; S02330.
DR HSP; P00802; 2SAK.
KW Hydrolase; Calcium; Plasminogen activation; Signal.
FT SIGNAL 1 27 POTENTIAL
FT CHAIN 28 163 STAPHYLOKINASE.
SQ SEQUENCE 163 AA; 18608 MW; AA2674FF75D365EC CRC64;

Query Match          97.2%; Score 695; DB 1; Length 163;
Best Local Similarity 97.8%; Pred. No. 2.6e-57;
Matches 133; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSFDKGYKKGDDASFEPTGYLMVNTGVDSKGNELLSPHYVEFPKPGTTLTKKEI 60
   |||||
Db 28 SSSFDKGYKKGDDASFEPTGYLMVNTGVDSKGNELLSPHYVEFPKPGTTLTKKEI 87
   |||||

QY 61 EYVVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSFPITEKGFVVPDLSEHI 120
   |||||
Db 88 EYVVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSFPITEKGFVVPDLSEHI 147
   |||||
```

```
QY 121 KNPGENLITKVIEKK 136
   |||||
Db 148 KNPGENLITKVIEKK 163
   |||||

RESULT 3
YB13_FOWPM
ID YB13_FOWPM STANDARD; PRT; 172 AA.
AC P14368;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL BAMHI-ORF13 PROTEIN (FRAGMENT).
OS Fowlpox virus (isolate HP-438[Munich]).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID=10263;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88229622; PubMed=2836548;
RA Tomley F., Binns M., Campbell J., Boursnell M.;
RT "Sequence analysis of an 11.2 kilobase, near-terminal, BamHI fragment
RT of fowlpox virus.";
RL J. Gen. Virol. 69:1025-1040(1988).
CC -!- SIMILARITY: CONTAINS AT LEAST 5 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D00295; BAA00209.1; -
DR PIR; D30087; WNVZTH.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 4.
DR SMART; SM00248; ANK; 3.
DR PROSITE; PS00088; ANK_REPEAT; 3.
DR PROSITE; PS0297; ANK_REPEAT_REGION; 1.
KW Hypothetical protein; Early protein; Repeat; ANK repeat.
FT NON_TER 1 1
FT REPEAT 7 36 ANK 1.
FT REPEAT 39 68 ANK 2.
FT REPEAT 73 105 ANK 3.
FT REPEAT 110 138 ANK 4.
FT REPEAT 143 172 ANK 5.
SQ SEQUENCE 172 AA; 19560 MW; 1929D02DC8C86457 CRC64;

Query Match          11.8%; Score 84.5; DB 1; Length 172;
Best Local Similarity 25.2%; Pred. No. 0.48;
Matches 31; Conservative 19; Mismatches 36; Indels 37; Gaps 5;

QY 43 HYVEFPIKPGTTLTKKEI-EYVVEWALDATA-----YKEFRV---VEL 81
   |||
Db 38 HYCLAPIHIAANRTSKIVKLLIEYGADINSESGANGKYPHYAMKVYDPPRLKIIVLL 97
   |||

QY 82 DPSAKIEVYYDKNKKKEETKSFPITEKGFVVPDLSEHIKNPGFNL-----ITKV 132
   |||
Db 98 DIGA-----DINKQSVLTNTSPLYETRFITDLDLYIIRGANINKEWVEVITRII 150
   |||

QY 133 IEK 135
   |||
Db 151 LER 153
   |||

RESULT 4
CADN_DROME
ID CADN_DROME STANDARD; PRT; 3097 AA.
AC O15943; Q9VJ7;
```

DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 CC NEURAL-CADHERIN PRECURSOR (CADHERIN-N PROTEIN) (DN-CADHERIN).  
 GN CADN OR CG7100.  
 OS Drosophila melanogaster (Fruit fly).  
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Head, and Embryo;  
 RX MEDLINE=97388431; PubMed=9247265;  
 RA Iwai Y., Usui T., Hirano S., Steward R., Takeichi M., Uemura T.;  
 RT "Axon patterning requires DN-cadherin, a novel neuronal adhesion  
 RL receptor, in the Drosophila embryonic CNS."; Neuron 19:77-89(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champe M., Pfeiffer B.D.,  
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berland B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster";  
 RL Science 287:2185-2195(2000).  
 RN [3]  
 RP INTERACTION WITH ARM.  
 RX MEDLINE=98298928; PubMed=9635189;  
 RA Loureiro J., Peifer M.;  
 RT "Roles of Armadillo, a Drosophila catenin, during central nervous  
 RT system development."; Curr. Biol. 8:622-632(1998).  
 RL CC  
 CC -!- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.  
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC  
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE  
 CC SORTING OF HETEROGENEOUS CELL TYPES. MAY ASSOCIATE WITH ARM NEURAL  
 CC ISOFORM AND PARTICIPATE IN THE TRANSMISSION OF DEVELOPMENTAL  
 CC INFORMATION.  
 CC  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).  
 CC -!- TISSUE SPECIFICITY: IN THE EMBRYO, THE PROTEIN FIRST APPEARS IN  
 CC THE MESODERM AT STAGE 9 AND IS PRESENT IN THE MYOBLASTS AND MUSCLE  
 CC FIBERS BY STAGE 12 AND STAGE 14, RESPECTIVELY. AT STAGE 12 THE  
 CC PROTEIN IS ALSO LOCATED IN THE AXONS OF THE ENTIRE CNS, BUT NOT IN  
 CC THE GLIAL CELLS. IN THIRD INSTAR LARVAE PROTEIN IS EXPRESSED IN  
 CC THE CNS NEUROPILE, PHOTORECEPTOR AXONS AND PRECURSORS OF ADULT  
 CC MUSCLES.  
 CC -!- SIMILARITY: CONTAINS 16 CADHERIN DOMAINS.  
 CC -!- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.  
 CC -!- SIMILARITY: CONTAINS 2 LAMININ G-LIKE DOMAINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AB002397; BAA22151.1; -;  
 CC EMBL; AE003656; AAF53635.1; -;  
 CC HSSP; P00740; 1IXA.  
 CC FlyBase; FBgn0015609; CadN.  
 CC InterPro; IPR002126; Cadherin.  
 CC InterPro; IPR000233; Cadherin\_C-term.  
 CC InterPro; IPR000561; EGF-like.  
 CC InterPro; IPR000742; EGF\_2.  
 CC InterPro; IPR001881; EGF\_Ca.  
 CC InterPro; IPR001791; Laminin\_G.  
 CC Pfam; PF000028; cadherin; 14.  
 CC Pfam; PF01049; Cadherin\_C-term; 1.  
 CC Pfam; PF00008; EGF; 3.  
 CC Pfam; PF00054; laminin\_G; 2.  
 CC PRINTS; PR00205; CADHERIN.  
 CC SMART; SM00112; CA; 16.  
 CC SMART; SM00179; EGF\_CA; 1.  
 CC SMART; SM00001; EGF\_like; 3.  
 CC SMART; SM00282; LamG; 2.  
 CC PROSITE; PS00232; CADHERIN\_1; 9.  
 CC PROSITE; PS0268; CADHERIN\_2; 16.  
 CC PROSITE; PS00022; EGF\_1; 3.  
 CC PROSITE; PS01186; EGF\_2; 3.  
 CC Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;  
 CC Signal; EGF-like domain.  
 CC SIGNAL; 1 36  
 CC PROPEP 37 ?  
 CC CHAIN ? 3097 NEURAL-CADHERIN.  
 CC DOMAIN ? 1454 EXTRACELLULAR (POTENTIAL).  
 CC TRANSMEM 1455 1475 POTENTIAL.  
 CC DOMAIN 1476 3097 CYTOPLASMIC (POTENTIAL).  
 CC DOMAIN 181 305 CADHERIN 1.  
 CC DOMAIN 430 543 CADHERIN 2.  
 CC DOMAIN 554 651 CADHERIN 3.  
 CC DOMAIN 660 756 CADHERIN 4.  
 CC DOMAIN 765 858 CADHERIN 5.  
 CC DOMAIN 867 968 CADHERIN 6.  
 CC DOMAIN 978 1078 CADHERIN 7.  
 CC DOMAIN 1087 1183 CADHERIN 8.  
 CC DOMAIN 1193 1299 CADHERIN 9.  
 CC DOMAIN 1307 1414 CADHERIN 10.  
 CC DOMAIN 1423 1534 CADHERIN 11.  
 CC DOMAIN 1523 1630 CADHERIN 12.  
 CC DOMAIN 1639 1742 CADHERIN 13.  
 CC DOMAIN 1749 1861 CADHERIN 14.  
 CC DOMAIN 1870 1966 CADHERIN 15.  
 CC DOMAIN 1974 2085 CADHERIN 16.  
 CC DOMAIN 2346 2377 EGF-LIKE 1.  
 CC DOMAIN 2407 2585 LAMININ G-LIKE 1.  
 CC DOMAIN 2592 2627 EGF-LIKE 2.  
 CC DOMAIN 2869 2902 EGF-LIKE 3.  
 CC DOMAIN 2661 2822 LAMININ G-LIKE 2.  
 CC DOMAIN 2346 2357 POTENTIAL.  
 CC DISULFID 2346

```
FT DISULFID 2351 2366 POTENTIAL.
FT DISULFID 2368 2377 POTENTIAL.
FT DISULFID 2592 2607 POTENTIAL.
FT DISULFID 2601 2616 POTENTIAL.
FT DISULFID 2618 2627 POTENTIAL.
FT DISULFID 2869 2880 POTENTIAL.
FT DISULFID 2874 2891 POTENTIAL.
FT DISULFID 2893 2902 POTENTIAL.
FT CARBOHYD 97 97 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 150 150 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 325 325 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 426 426 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 930 930 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1266 1266 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 1425 1425 E -> K (IN ALLELE CADN-M12; MUSCLE DEFECTS).
FT CONFLICT 1342 1342 P -> A (IN REF. 1).
FT CONFLICT 2786 2786 S -> T (IN REF. 1).
SQ SEQUENCE 3097 AA; 347201 MW; 082242F28D9B5CC3 CRC64;

Query Match 11.3%; Score 80.5; DB 1; Length 3097;
Best Local Similarity 27.6%; Pred. No. 37;
Matches 45; Conservative 16; Mismatches 55; Indels 47; Gaps 9;

QY 13 DDASYFEPTGPLYMNVNVTGVDKGNELSP-----HYVEFPKPGTTLTKK----- 59
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 731 EDGGFSTVSDLTIRVTDVNDNAPKFPDLYQAHHNVDIEDPLGTSILRVKAMDSGSSN 789
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 60 --TEYVV---EWALDA-----TAYKEFRVVELD-----PSAKIEVTYDK 94
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 790 AEIYLVSDHFAVDNSGIIVNNKQLDADNNNAYEIVTAKDKGEPKSGVATRVYTK 849
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 95 NKKKEETKSPITEKGFVVDLSEHKINPGFN-LITKVVIEKK 136
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 850 NKDEE---PKFSQQVYTNVDE---NAGPNTLVTTVVASDK 885
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
YF02_METJA STANDARD; PRT; 324 AA.
AC Q58897;
AT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN MJ1502.
GN MJ1502.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Cocayne J.D.,
RA Kervatage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Otterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii."
RL Science 273:1058-1073(1996).
CC -!- SIMILARITY: BELONGS TO THE ATSA / ELAC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y11996; CAAY2721.1;
CC InterPro; IPR000504; RRM.
CC Pfam; PF00076; rrm; 1.
CC SMART; SM00360; RRM; 1.
CC PROSITE; PS00102; RRM; 1.
CC PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
KW Initiation factor; Protein biosynthesis; RNA-binding.
FT DOMAIN 60 147 RNA-BINDING (RRM).
SQ SEQUENCE 719 AA; 82524 MW; BBFA223B7DA52BB2 CRC64;
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67591; AAB99515.1;
CC TIGR; MJ1502;
CC InterPro; IPR001279; Beta_lactam_met.
CC Pfam; PF00753; lactamase_B; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 324 AA; 36893 MW; A3DF62E3BAAEEFA4 CRC64;

Query Match 11.0%; Score 78.5; DB 1; Length 324;
Best Local Similarity 24.2%; Pred. No. 3.7;
Matches 39; Conservative 18; Mismatches 57; Indels 47; Gaps 7;

QY 7 GYKKGDDASYFEPTGPLYMNVNVTGVDKGNEL--LSPHYVEFPKPGTTLTKKIEYVE 65
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 87 GFGREKELKIFGEGTKEIIE-----NSLKLGHYIEFPKIVYIEYTKPEITTYKE 138
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 66 WALDATAY-----KEFRVVVELD--PSAKIEVTYDKNKK----- 97
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 139 ENVEIIAYPTEHGIPSVAYIFKEIKKPRLDIERAKKLGKIGDPDKLKNGEAVKNIYGE 198
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 98 --KEE-----TKSPITEKGFVVP--DLSEHIKNPGFNLI 128
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 199 IIRPEVLLPPKKGKCLAYSGDTLPLEDFGKYLKELGCDVL 239
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
IF39_TOBAC STANDARD; PRT; 719 AA.
ID IF39_TOBAC
AC P56821;
AT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE EUKARYOTIC TRANSLATION INITIATION FACTOR 3 SUBUNIT 9 (EIF-3 ETA) (EIF3 P110).
DE P110).
GN PRT1.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. BRIGHT YELLOW 2;
RA Shen W.H., Gigot C.;
RT "Characterization of prt1, a gene encoding for one of the subunits of
RT the translation initiation factor 3 (eif3), from Nicotiana tabacum.";
RL Plant Sci. 143:45-54(1999).
CC -!- FUNCTION: BINDS TO THE 40S RIBOSOME AND PROMOTES THE BINDING OF
CC METHIONYL-TRNAI AND MRNA.
CC -!- SUBUNIT: EIF-3 IS COMPOSED OF AT LEAST 10 DIFFERENT SUBUNITS (BY
CC SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y11996; CAAY2721.1;
CC InterPro; IPR000504; RRM.
CC Pfam; PF00076; rrm; 1.
CC SMART; SM00360; RRM; 1.
CC PROSITE; PS00102; RRM; 1.
CC PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
KW Initiation factor; Protein biosynthesis; RNA-binding.
FT DOMAIN 60 147 RNA-BINDING (RRM).
SQ SEQUENCE 719 AA; 82524 MW; BBFA223B7DA52BB2 CRC64;
```

```

KW Isomerase; Topoisomerase; DNA-binding.
FT ACT_SITE 117 117 5'-ENDS OF THE DNA VIA A PHOSPHO-TYROSYL
SQ SEQUENCE 442 AA; 50493 MW; 51960AEE6F75AF7 CRC64;

Query Match 10.8%; Score 77.5; DB 1; Length 442;
Best Local Similarity 25.5%; Pred. No. 6.7; Indels 41; Caps 8;
Matches 36; Conservative 21; Mismatches 43;

QY 4 FDGKYKKGDASYF-----EPTGYLMVNVTGVDKSGN-----ELLSPH 43
DB 543 FDVELETWASAEHFWATDVEWDPTGYVATSVTSHEMENGNIWSFGKLLYRLKDH 602

QY 44 YVEF---PIKPGTTLTKIEIYVVEWALDATAT-KEFRVVVLDPSAKIEVYIYVYDKNKK-K 98
DB 603 FFQYLWRP-RPPSFLSKREE---ETAKNLKRYSKKYAEADQDVLSQLSDQDRKKKLLK 658

QY 99 EETKSF 104
DB 659 EEWEAW 664

RESULT 7
TOP5_BPT4 STANDARD; PRT; 442 AA.
AC P07065;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE DNA TOPOISOMERASE MEDIUM SUBUNIT (EC 5.99.1.3) (PROTEIN GP52).
GN 52.
OS Bacteriophage T4.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae;
OC T4-like phages.
OX NCBI_TaxID=10665;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87016377; PubMed=3020513;
RA Huang W.M.;
RT "The 52-protein subunit of T4 DNA topoisomerase is homologous to the
RL gyra-protein of gyrase."
RL Nucleic Acids Res. 14:7379-7390(1986).
RN [2]
RP REVISIONS.
RA Huang W.M.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-46 FROM N.A.
RX MEDLINE=88172481; PubMed=3280805;
RA Chapman D., Morad I., Kaufmann G., Gait M.J., Jorissen L., Snyder L.;
RT "Nucleotide and deduced amino acid sequence of stp: the bacteriophage
RL T4 anticodon nuclease gene."
RL J. Mol. Biol. 199:373-377(1988).
CC -!- FUNCTION: CONTROL OF TOPOLOGICAL STATES OF DNA BY TRANSIENT
CC BREAKAGE AND SUBSEQUENT REJOINING OF DNA STRANDS. T4 TOPOISOMERASE
CC MAKES DOUBLE-STRAND BREAKS.
CC -!- CATALYTIC ACTIVITY: ATP-DEPENDENT BREAKAGE, PASSAGE AND REJOINING
CC OF DOUBLE-STRANDED DNA.
CC -!- SUBUNIT: THIS IS ONE OF THE 3 SUBUNITS OF T4-DNA TOPOISOMERASE.
CC -!- SIMILARITY: TO THE PROKARYOTIC GYRASE SUBUNIT A AND TO EUKARYOTIC
CC TOPOISOMERASE II.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X04376; CAA27959.1; ..
CC PIR: B24705; ITBPT4.
CC DR PIR: S01872; S01872.
CC DR InterPro: IPR002205; DNA_topoisoiV.
CC DR Pfam: PF00521; DNA_topoisoiV; 1.
CC DR SMART: SM00434; TOP4c; 1.
CC -----

KW Isomerase; Topoisomerase; DNA-binding.
FT ACT_SITE 117 117 5'-ENDS OF THE DNA VIA A PHOSPHO-TYROSYL
SQ SEQUENCE 442 AA; 50493 MW; 51960AEE6F75AF7 CRC64;

Query Match 10.8%; Score 77.5; DB 1; Length 442;
Best Local Similarity 25.5%; Pred. No. 6.7; Indels 41; Caps 8;
Matches 36; Conservative 21; Mismatches 43;

QY 16 SYEPTGPLYLMVN-VTGVDKSGNELLSPHYVEFPK-----PGTTLTKIEIYVVEWAL 68
DB 148 AFVLPITPTVLLGVSGIATGYATILPHSVSSVKRAVLQALQGLQKVKTPKVE----- 200

QY 69 DATAYKEFR--VVELDPSAKIEVY-----YDNKKKEETKSF-PIEKGK 111
DB 201 ----FPEFRGEVVEIDQVEIRGTFTSRTHMHITEIPYKYDRETVYVSKILDPLENKGF 256

QY 112 VVPD--LSEHIKNPGENLITK 130
DB 257 ITWDDACGEH-----GFGFKVK 273

RESULT 8
PRLR_MOUSE STANDARD; PRT; 608 AA.
ID PRLR_MOUSE
AC Q08501; Q62099; P15213; P15212;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROLACTIN RECEPTOR PRECURSOR (PRL-R).
GN PRLR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (FORM PRL-R3).
RC STRAIN=C3H; TISSUE=Mammary gland;
RX MEDLINE=94085788; PubMed=8262385;
RA Moore R.C., Oka T.;
RT "Cloning and sequencing of the cDNA encoding the murine mammary gland
RL long-form prolactin receptor."
RL Gene 134:263-265(1993).
RN [2]
RP SEQUENCE FROM N.A. (FORM PRL-R3).
RC STRAIN=SWISS WEBSTER; TISSUE=Liver;
RX MEDLINE=93307149; PubMed=8319571;
RA Clarke D.L., Linzer D.I.H.;
RT "Changes in prolactin receptor expression during pregnancy in the
RL mouse ovary."
RL Endocrinology 133:224-232(1993).
RN [3]
RP SEQUENCE FROM N.A. (FORM PRL-R3).
RA Sasaki M.;
RL Submitted (JUL-1992) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (FORM PRL-R3).
RC STRAIN=BALB/C; TISSUE=Mammary gland;
RA Edery M., Pezet A., Nandi S., Kelly P.A.;
RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (FORMS PRL-R2 AND PRL-R1).
RC STRAIN=SWISS WEBSTER; TISSUE=Liver;
RX MEDLINE=89261824; PubMed=2725531;
RA Davis J.A., Linzer D.I.H.;
RT "Expression of multiple forms of the prolactin receptor in mouse
RL liver."
RL Mol. Endocrinol. 3:674-680(1989).
CC -!- FUNCTION: THIS IS A RECEPTOR FOR THE ANTERIOR PITUITARY HORMONE
CC PROLACTIN, AS WELL AS PLACENTAL LACTOGEN I AND II.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; PRL-R1, PRL-R2 AND PRL-R3 (SHOWN
```

```
CC HERE): ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.
CC -1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L13593; AAC37641.1; -
DR EMBL; L14811; AAA02686.1; -
DR EMBL; D10214; BAA01066.1; -
DR EMBL; X73372; AAA51789.1; -
DR EMBL; M22959; AAA39977.1; -
DR EMBL; M22958; AAA39976.1; -
DR PIR; J0671; J0671.
DR HSP; P16471; IBP3.
DR MGD; MGI:97763; PrIr.
DR InterPro; IPR002996; CR1A.
DR InterPro; IPR001777; FN III.
DR InterPro; IPR003528; Hematopo_receptor_L_F1.
DR Pfam; PF00041; fn3; 2.
DR SMART; SM00060; FN3; 1.
DR PROSITE; PS01352; HEMATOPO_REC_L_F1; 1.
DR Receptor; Transmembrane; Glycoprotein; Signal; Repeat;
KW Alternative splicing.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 608 PROLACTIN RECEPTOR.
FT DOMAIN 20 229 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 230 253 BY SIMILARITY.
FT DOMAIN 254 608 CYTOPLASMIC (BY SIMILARITY).
FT DOMAIN 20 117 FIBRONECTIN TYPE-III 1.
FT DOMAIN 119 222 FIBRONECTIN TYPE-III 2.
FT DISULFID 31 41 BY SIMILARITY.
FT DISULFID 70 81 BY SIMILARITY.
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 281 292 KGSEELLSALG -> VHNKEQLENVY (IN ISOFORM
FT PRL-R2).
FT VARSPLIC 293 608 MISSING (IN ISOFORM PRL-R2).
FT VARSPLIC 281 303 KGSEELLSALGQDFPPTSDCE -> LMCSILQLSLVKI
FT PTFEFLCDL (IN ISOFORM PRL-R1).
FT VARSPLIC 304 608 MISSING (IN ISOFORM PRL-R1).
FT CONFLICT 558 558 L -> F (IN REF. 2).
SQ SEQUENCE 608 AA; 68240 MW; B8CE202B2EFC9FC6 CRC64;

Query Match 10.8%; Score 77; DB 1; Length 608;
Best Local Similarity 23.9%; Pred. No. 11;
Matches 34; Conservative 17; Mismatches 43; Indels 48; Gaps 7;

QY 4 FDGKGYKGGDASY-----PEPTGPY--LMVNTGVDSKGNELLSP-HYVEFP---I 49
|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 274 FDTLLKSGKSELLSALGQDFPPTSDCEDLLVEFLVEDNDERLMPHSKEYEPQGV 333
|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 50 KPGTTLTKKIEYVVEWALDATAYKEPRVVELDPSSAKIEVYDK-----NKKKEETKSPF 105
|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 334 KP-----THLPDSDSGHGSYDSSLSEKCEEPQAYP 366
|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 106 ITEKGFVVPDLSHKINPGPNL 127
|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 367 ---PAFHPIETETKPEANI 385
|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
VLI_BPV4
ID VLI_BPV4 STANDARD; PRT; 506 AA.
AC P08341;
DT 01-AUG-1988 (Rel. 08, Created)
```

```
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MAJOR CAPSID PROTEIN L1.
GN L1.
OS Bovine papillomavirus type 4.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87282264; PubMed=3039043;
RA Patel K.R., Smith K.T., Campo M.S.;
RT "The nucleotide sequence and genome organization of bovine
RL papillomavirus type 4."
RL J. Gen. Virol. 68:2117-2128(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X05817; -; NOT_ANNOTATED_CDS.
DR EMBL; D00146; BAA00101.1; -
DR PIR; B26214; PIWLB4.
DR InterPro; IPR002210; PV_capsid_L1.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVCAPSIDL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
KW Coat protein; Late protein.
SQ SEQUENCE 506 AA; 57902 MW; FLC5DFDB54FA681E CRC64;

Query Match 10.7%; Score 76.5; DB 1; Length 506;
Best Local Similarity 28.7%; Pred. No. 9.7;
Matches 29; Conservative 19; Mismatches 36; Indels 17; Gaps 7;

QY 29 TVGVDS-KG-NELLSPHYVEFPKPGTTLTKKIEYV-----EWALDATAYKEFRVVELD 82
|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 331 VTAVDSTRGNFSVHTDPEVKPQETATATKFKHYLRHVEW--DLSLMQLCIVNLT 388
|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 83 PSKIEVYTYDKNKKKEETKSPFITEKGFVVP--DLSEHIK 121
|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 389 P-----ESIAYLHNWNSIENNNL---GFTQPPNDIEDHYR 422
|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
LXBL_PHOLU
ID LXBL_PHOLU STANDARD; PRT; 327 AA.
AC P19840;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE ALKANAL MONOOXYGENASE BETA CHAIN (EC 1.14.14.3) (BACTERIAL LUCIFERASE
DE BETA CHAIN).
GN LUXB.
OS Photorhabdus luminescens (Xenorhabdus luminescens).
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Photorhabdus.
OX NCBI_TaxID=29488;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 29999;
RX MEDLINE=90375532; PubMed=2204626;
RA Szitner R., Weighen E.;
RT "Nucleotide sequence, expression, and properties of luciferase coded
RL by lux genes from a terrestrial bacterium."
RL J. Biol. Chem. 265:16581-16587(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 29999;
```



RX MEDLINE=92355513; PubMed=1644764;  
RA Meighen E.A., Szittner R.B.;  
RT "Multiple repetitive elements and organization of the lux operons of  
luminescent terrestrial bacteria";  
RL J. Bacteriol. 174:5371-5381(1992).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HM;  
RX MEDLINE=90343746; PubMed=2383248;  
RA Johnston T.C., Rucker E.B., Cochrum L., Hruska K.S., Vandegrift V.;  
RT "The nucleotide sequence of the luxA and luxB genes of Xenorhabdus  
luminescens HM and a comparison of the amino acid sequences of  
luciferases from four species of bioluminescent bacteria";  
RL Biochem. Biophys. Res. Commun. 170:407-415(1990).  
CC -|- FUNCTION: LIGHT-EMITTING REACTION IN LUMINOUS BACTERIA. THE  
SPECIFIC ROLE OF THE BETA SUBUNIT IS UNKNOWN, BUT IT IS ABSOLUTELY  
REQUIRED FOR BIOLUMINESCENCE ACTIVITY.  
CC -|- CATALYTIC ACTIVITY: FMNH(2) + RCHO + O(2) = FMN + RCOOH + H(2)O  
+ LIGHT.  
CC -|- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.  
CC  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: M57416; AAA27624.1; -  
DR EMBL: M90093; AAA27620.1; -  
DR EMBL: M55977; AAA27627.1; -  
DR PIR: C37898; C37898.  
DR HSP: P07739; 1XKJ.  
DR InterPro: IPR002103; Bac\_luciferase.  
DR Pfam: PF00296; bac\_luciferase; 1.  
DR PRINTS: P00089; LUCIFERASE.  
DR PROSITE: PS00494; BACTERIAL\_LUCIFERASE; 1.  
KW Photoprotein; Luminescence; Oxidoreductase; Monooxygenase;  
FT Flavoprotein; FMN.  
FT VARIANT 85 A -> R (IN STRAIN HM).  
SQ SEQUENCE 327 AA; 37595 MW; 1905AF01658BE56E CRC64;

Query Match 10.6%; Score 75.5; DB 1; Length 327;  
Best Local Similarity 25.4%; Pred. No. 7.1;  
Matches 33; Conservative 18; Mismatches 44; Indels 35; Gaps 7;  
QY 18 FEPTGPLYMNVTVGDSKGNELLSPHYVEFPKPGTTL-----TKETIYVVEWALDA 70  
Db 163 YTPGGRKKYVTA-----SHHIVWAAKKGIPLIFKWDSDNDVRYEAERYKAVA 212  
QY 71 TAYKEFRVVELDPSAKIEVY-YDRNKKKEETKSFPITEKGFVVPDLSHII-----KNPGF 125  
Db 213 DKY-DVDLSDHQLMLVYNEDSNKAKQETRAF-----ISDYVLEHMHENF 260  
QY 126 -NLITKVVIE 134  
Db 261 ENKLEETIAE 270  
RESULT 11  
NAE2\_THEME  
ID NAE2\_THEME STANDARD; PRT; 576 AA.  
AC Q9X0Y0;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE PROBABLE GLUTAMINE-DEPENDENT NAD(+) SYNTHETASE (EC 6.3.5.1) (NAD(+)  
DE SYNTHASE [GLUTAMINE-HYDROLYSING]).  
GN NAE2 OR TM1253.  
OS Thermotoga maritima.  
OC Bacteria; Thermotogales; Thermotoga.

OX NCBI\_TaxID=23336;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MSB8 / DSM 3109;  
RX MEDLINE=99287316; PubMed=10360571;  
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,  
Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,  
McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,  
Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,  
Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,  
Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;  
RT "Evidence for lateral gene transfer between Archaea and Bacteria from  
genome sequence of Thermotoga maritima";  
RL Nature 399:323-329(1999).  
CC -|- FUNCTION: CAN USE BOTH GLUTAMINE OR AMMONIA AS A NITROGEN  
SOURCE (BY SIMILARITY).  
CC -|- CATALYTIC ACTIVITY: ATP + DEAMIDO-NAD(+) + L-GLUTAMINE + H(2)O =  
AMP + DIPHOSPHATE + NAD(+) + L-GLUTAMATE.  
CC -|- PATHWAY: DE NOVO BIOSYNTHESIS OF NAD.  
CC -|- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE NAD  
SYNTHETASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AF001780; NAD36328.1; -  
DR TIGR: TM1253; -  
DR InterPro: IPR003010; CN\_hydrolase.  
DR InterPro: IPR003694; NAD\_synthase.  
DR InterPro: IPR000132; Nitril\_cyn\_hydrolase.  
DR Pfam: PF00795; CN\_hydrolase; 1.  
DR Pfam: PF02540; NAD\_synthase; 1.  
DR PROSITE: PS00920; NITRIL\_CHT\_1; UNKNOWN\_1.  
KW Ligase; NAD; ATP-binding; Complete proteome.  
FT DOMAIN 292 576 LIGASE.  
FT NP\_BIND 321 328 ATP (BY SIMILARITY).  
FT ACT\_SITE 323 323 BY SIMILARITY.  
SQ SEQUENCE 576 AA; 64919 MW; 208FDC77964C957F CRC64;  
Query Match 10.5%; Score 75; DB 1; Length 576;  
Best Local Similarity 18.1%; Pred. No. 16;  
Matches 39; Conservative 31; Mismatches 54; Indels 92; Gaps 7;  
QY 7 GYKK-----GDDASYFEPTGPLYMNV-----TGV 32  
Db 108 GYRKISLPNYGVFDERRYPKGEELLVKIGIKVGVITCEDINWVPSASLSLGEV 167  
QY 33 DSKGNELLSPHYVEFPKPGTTLTKETIYVVEWA-----LDATA- 72  
Db 168 HLIALNSASPHYGKPKVLRKDYLSMKAYDHYVAMAYCNVVGQDELVDFDGGSMVDASGE 227  
QY 73 -----YKEFRVVELDPSAKIEVY-YDRNKKKEETKSFPI--TEKG----- 110  
Db 228 VINYGKLFEEIITVDLDLDENLRVSLVDPRRRYMKQTPYKVTVEAGNLRKSGHPEV 287  
QY 111 -----FVVPDLSEHIKNPGFNLTKKVI 133  
Db 288 VNPLPVREEMFRALITGLRDYVRKNGFE---KVVI 320  
RESULT 12  
FAS\_CHICK  
ID FAS\_CHICK STANDARD; PRT; 2511 AA.  
AC P12276;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE FATTY ACID SYNTHASE (EC 2.3.1.85) [INCLUDES: EC 2.3.1.38; EC 2.3.1.39;  
DN EC 2.3.1.41; EC 1.1.1.100; EC 4.2.1.61; EC 1.3.1.10; EC 3.1.2.14].  
GN FASN OR FAS.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OC NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A. AND SEQUENCE OF 1-11.  
RC STRAIN=WHITE LEGHORN; TISSUE=Liver;  
RX MEDLINE=89282777; PubMed=2734291;  
RA Huang W.-Y., Chirala S.S., Wakil S.J.;  
RT "Amino-terminal blocking group and sequence of the animal fatty acid  
synthase.";  
RL Arch. Biochem. Biophys. 314:45-49(1994).  
RN [2]  
RP SEQUENCE OF 75-1775 FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=89282777; PubMed=2734291;  
RA Holzer K.P., Liu W., Hammes G.G.;  
RT "Molecular cloning and sequencing of chicken liver fatty acid  
synthase cDNA.";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:4387-4391(1989).  
RN [3]  
RP SEQUENCE OF 1568-2512 FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE=89139426; PubMed=2917973;  
RA Chirala S.S., Kasturi R., Pazirandeh M., Stelow D.T., Huang W.-Y.,  
RA Wakil S.J.;  
RT "A novel cDNA extension procedure. Isolation of chicken fatty acid  
synthase cDNA clones.";  
RL J. Biol. Chem. 264:3750-3757(1989).  
RN [4]  
RP SEQUENCE OF 1752-2512 FROM N.A.  
RX MEDLINE=88320436; PubMed=2842766;  
RA Yuan Z., Liu W., Hammes G.G.;  
RT "Molecular cloning and sequencing of DNA complementary to chicken  
liver fatty acid synthase mRNA.";  
RL Proc. Natl. Acad. Sci. U.S.A. 85:6328-6331(1988).  
RN [5]  
RP SEQUENCE OF 2202-2512 FROM N.A.  
RX MEDLINE=89088152; PubMed=3207710;  
RA Kasturi R., Chirala S.S., Pazirandeh M., Wakil S.J.;  
RT "Characterization of a genomic and cDNA clone coding for the  
thioesterase domain and 3' noncoding region of the chicken liver  
fatty acid synthase gene.";  
RL Biochemistry 27:7778-7785(1988).  
RN [6]  
RP SEQUENCE OF 2121-2209.  
RX MEDLINE=89192401; PubMed=2648999;  
RA Huang W.-Y., Stoops J.K., Wakil S.J.;  
RT "Complete amino acid sequence of chicken liver acyl carrier protein  
derived from the fatty acid synthase.";  
RL Arch. Biochem. Biophys. 270:92-98(1989).  
RN [7]  
RP SEQUENCE OF 2209-2508.  
RC STRAIN=WHITE LEGHORN;  
RX MEDLINE=89088151; PubMed=3207709;  
RA Yang C.-Y., Huang W.-Y., Chirala S.S., Wakil S.J.;  
RT "Complete amino acid sequence of the thioesterase domain of chicken  
liver fatty acid synthase.";  
RL Biochemistry 27:7773-7777(1988).  
RN [8]  
RP SEQUENCE OF 667-674 AND 1698-1709.  
RX MEDLINE=89323081; PubMed=2751995;  
RA Chang S.I., Hammes G.G.;  
RT "Amino acid sequences of pyridoxal 5'-phosphate binding sites and  
fluorescence resonance energy transfer in chicken liver fatty acid  
synthase.";  
RL Biochemistry 28:3781-3788(1989).  
CC -1- FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF  
CC LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.  
CC THIS MULTIFUNCTIONAL PROTEIN HAS 7 CATALYTIC ACTIVITIES AND AN

ACYL CARRIER PROTEIN.  
-1- CATALYTIC ACTIVITY: ACETYL-COA + N MALONYL-COA + 2N NADPH =  
LONG-CHAIN FATTY ACID + (N+1) COA + N CO(2) + 2N NADP(+).  
-1- CATALYTIC ACTIVITY: ACETYL-COA + [ACYL-CARRIER PROTEIN] = COA  
+ ACETYL-[ACYL-CARRIER PROTEIN].  
-1- CATALYTIC ACTIVITY: MALONYL-COA + [ACYL-CARRIER PROTEIN] = COA  
+ MALONYL-[ACYL-CARRIER PROTEIN].  
-1- CATALYTIC ACTIVITY: ACYL-[ACYL-CARRIER PROTEIN] + MALONYL-[ACYL-  
CARRIER PROTEIN] = 3-OXOACYL-[ACYL-CARRIER PROTEIN] + CO(2) +  
[ACYL-CARRIER PROTEIN].  
-1- CATALYTIC ACTIVITY: (3R)-3-HYDROXYACYL-[ACYL-CARRIER PROTEIN] +  
NADP(+) = 3-OXOACYL-[ACYL-CARRIER PROTEIN] + NADPH.  
-1- CATALYTIC ACTIVITY: (3R)-3-HYDROXYPALMITOYL-[ACYL-CARRIER PROTEIN]  
= 2-HEXADECENOYL-[ACYL-CARRIER PROTEIN] + H(2)O.  
-1- CATALYTIC ACTIVITY: ACYL-[ACYL-CARRIER PROTEIN] + NADP(+) = 2,3-  
DEHYDROACYL-[ACYL-CARRIER PROTEIN] + NADPH.  
-1- CATALYTIC ACTIVITY: OLEOYL-[ACYL-CARRIER PROTEIN] + H(2)O =  
ACYL-CARRIER PROTEIN + OLEATE.  
-1- SUBUNIT: HOMODIMER, WHICH IS ARRANGED IN A HEAD TO TAIL FASHION.  
-1- CAUTION: REF.3 AND REF.5 SEQUENCES DIFFER FROM THAT SHOWN FROM  
POSITION 2351 ONWARDS DUE TO A FRAMESHIFT.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement. See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
-----  
EMBL: J04485; AAB46389.1; -  
DR EMBL: J03860; AAA48767.1; -  
DR EMBL: J02839; AAC82106.1; ALT\_SEQ.  
DR PIR: A33918; XYCHFA.  
DR PIR: A32015; A32015.  
DR HSP: P02901; LACP.  
DR InterPro: IPR001227; Acyltransf\_domain.  
DR InterPro: IPR002085; Adh\_zinc.  
DR InterPro: IPR000794; Ketoacyl-synt.  
DR InterPro: IPR003880; Phosphopant\_attach.  
DR InterPro: IPR001031; Thioesterase.  
DR Pfam: PF00698; Acyl\_transf; 1.  
DR Pfam: PF00107; adh\_zinc; 1.  
DR Pfam: PF00109; ketoacyl-synt; 1.  
DR Pfam: PF00550; pp-binding; 1.  
DR Pfam: PF00975; Thioesterase; 1.  
DR PROSITE: PS00012; PHOSPHOPANTETHEINE; 1.  
DR PROSITE: PS0075; ACP\_DOMAIN; 1.  
DR PROSITE: PS50075; ACP\_DOMAIN; 1.  
KW Fatty acid biosynthesis; Multifunctional enzyme; Phosphopantetheine;  
KW Transferase; Hydrolase; Oxidoreductase; Lyase; NADP; Acetylation;  
FT Alternative splicing.  
FT INIT\_MET 0 0  
FT DOMAIN 1 7411 BETA-KETOACYL SYNTHASE.  
FT DOMAIN 427 815 ACYL AND MALONYL TRANSFERASES.  
FT DOMAIN 1638 1866 ENOYL REDUCTASE.  
FT DOMAIN 1867 2119 BETA-KETOACYL REDUCTASE.  
FT DOMAIN 2124 2180 ACYL CARRIER (ACP).  
FT DOMAIN 2209 2511 THIOESTERASE.  
FT MOD\_RES 1 1 ACETYLATION.  
FT ACT\_SITE 160 160 BETA-KETOACYL SYNTHASE (BY SIMILARITY).  
FT ACT\_SITE 579 579 ACYL/MALONYL TRANSFERASES (BY  
SIMILARITY).  
FT ACT\_SITE 877 877 BETA-HYDROXYACYL DEHYDRATASE (BY  
SIMILARITY).  
FT NP\_BIND 1674 1691 NADP (ER).  
FT BINDING 1707 1707 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
FT NP\_BIND 1888 1903 NADP (KR).  
FT BINDING 2157 2157 PHOSPHOPANTETHEINE (BY SIMILARITY).  
FT ACT\_SITE 2308 2308 THIOESTERASE (BY SIMILARITY).  
FT ACT\_SITE 2481 2481 THIOESTERASE (BY SIMILARITY).  
FT VARSPIC 2348 2348 T -> TQCFSESLF (IN ISOFORM 1).  
FT CONFLICT 77 78 QL -> PV (IN REF. 2).

```
FT CONFLICT 116 116 L -> A (IN REF. 2).
FT CONFLICT 675 675 R -> S (IN REF. 2).
FT CONFLICT 1169 1169 K -> N (IN REF. 2).
FT CONFLICT 1178 1178 A -> T (IN REF. 2).
FT CONFLICT 1191 1191 R -> H (IN REF. 2).
FT CONFLICT 1198 1198 P -> L (IN REF. 2).
FT CONFLICT 1286 1287 DN -> ND (IN REF. 2).
FT CONFLICT 1372 1372 K -> E (IN REF. 2).
FT CONFLICT 1533 1533 C -> Y (IN REF. 2).
FT CONFLICT 1577 1577 W -> R (IN REF. 2).
FT CONFLICT 1685 1696 QAAIAIALSMGC -> ASSHCHRLHGLA
      (IN REF. 2; AAA48767).
FT CONFLICT 1732 1732 O -> E (IN REF. 2).
FT CONFLICT 1745 1745 S -> N (IN REF. 2).
FT CONFLICT 2511 AA; 274648 MW; 622039DAC8315D3F CRC64;
SQ SEQUENCE 2511 AA; 274648 MW; 622039DAC8315D3F CRC64;

Query Match 10.4%; Score 74.5; DB 1; Length 2511;
Best Local Similarity 25.5%; Pred. No. 1e+02;
Matches 26; Conservative 13; Mismatches 24; Indels 39; Gaps 5;

QY 27 VNVGTGDSKGNELLSHYVEFPKPGTTLTKKIEYVVEWALDATAKFRVVELDPSAK 86
DB 801 IHLTGINVLGNLFPB--VEYVPVGTPL-----ISPYIKW-----DHSQD 839

QY 87 IEVYYDKNKKKEETKSPITEKGF-----VVPDLSEH 119
DB 840 WDV-----PKAEDFSGSKGSASVYINIDVSPDSPDH 872

RESULT 13
MOAA_METJA
ID MOAA_METJA STANDARD; PRT; 298 AA.
AC Q58234;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PUTATIVE MOLYBDOPTERIN COFACTOR SYNTHESIS PROTEIN A.
GN MOAA OR MJ0824.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -!- FUNCTION: INVOLVED IN THE BIOSYNTHESIS OF MOLYBDOPTERIN PRECURSOR
CC Z FROM GUANOSINE (BY SIMILARITY).
CC -!- PATHWAY: MOLYBDENUM COFACTOR BIOSYNTHESIS (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE MOAA / NIFB / PQOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67526; AAB98823.1;
DR TIGR; MJ0824;

```

```
DR InterPro; IPR000385; MoAA_NifB_PqqP.
DR Pfam; PF01444; MoAA_NifB_PqqP; 1.
DR PROSITE; PS01305; MOAA_NIFB_PQOE; 1.
KW Molybdenum cofactor biosynthesis; Iron-sulfur; Complete proteome.
FT METAL 20 20 IRON-SULFUR (POTENTIAL).
FT METAL 24 24 IRON-SULFUR (POTENTIAL).
FT METAL 27 27 IRON-SULFUR (POTENTIAL).
SQ SEQUENCE 298 AA; 34652 MW; EA92B9A0A75BFD55 CRC64;

Query Match 10.3%; Score 74; DB 1; Length 298;
Best Local Similarity 21.4%; Pred. No. 8.7;
Matches 21; Conservative 19; Mismatches 42; Indels 16; Gaps 1;

QY 31 GVDKGNELLSHYVEFPKPGTTLTKKIEYVVEWALDATAKFRVVELDPSAKIEVT 90
DB 31 GHDSNDRYMTPEEIGIIARTKTKFGVKKI-----KISGGEPLLRKQVC 74

QY 91 YVDKNNKKKEETKSPITEKGFVVPDLSEHIKNPGENLI 128
DB 75 EIIENIKDERIKDISLTNGILLNLAELKLDAGLNRV 112

RESULT 14
PRLR_RAT
ID PRLR_RAT STANDARD; PRT; 610 AA.
AC P05710; Q63451; Q63723; Q62832; Q64274; Q63479;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PROLACTIN RECEPTOR PRECURSOR (PRL-R) (LACTOGEN RECEPTOR).
GN PRLR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91155946; PubMed=2293022;
RA Shiota M., Banville D., Ali S., Jolicoeur C., Boutin J.M.,
RA Edery M., Djiane J., Kelly P.A.;
RT "Expression of two forms of prolactin receptor in rat ovary and
RT liver."
RL Mol. Endocrinol. 4:1136-1143(1990).
RN [2]
RP SEQUENCE FROM N.A. (LONG FORM AND SHORT FORM).
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-Ovary;
RX MEDLINE=90241201; PubMed=2159291;
RA Zhang R., Buczko E., Tsai-Morris C.H., Hu Z.Z., Dufau M.L.;
RT "Isolation and characterization of two novel rat ovarian lactogen
RT receptor cDNA species."
RL Biochem. Biophys. Res. Commun. 168:415-422(1990).
RN [3]
RP SEQUENCE OF 281-610 FROM N.A.
RA Banville D., Stocco R., Murthy K.K., Boie Y., Kelly P.A.;
RL Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE (MEDIUM FORM).
RC TISSUE=Liver;
RX MEDLINE=88165059; PubMed=2832068;
RA Boutin J.-M., Jolicoeur C., Okamura H., Gagnon J., Edery M.,
RA Shiota M., Banville D., Dusanter-Fourt I., Djiane J., Kelly P.A.;
RT "Cloning and expression of the rat prolactin receptor, a member of
RT the growth hormone/prolactin receptor gene family."
RL Cell 53:69-77(1988).
RN [5]
RP SEQUENCE FROM N.A. (FORM NB2).
RC TISSUE=Lymphoma;
RX MEDLINE=92041834; PubMed=1718958;
RA Ali S., Pelligrini I., Kelly P.A.;
RT "A prolactin-dependent immune cell line (Nb2) expresses a mutant form
RT of prolactin receptor."
RL J. Biol. Chem. 266:20110-20117(1991).
```

RN SEQUENCE FROM N.A. (FORM NB2).  
 RP MEDLINE=95014432; PubMed=7929319;  
 RA O'Neal K.D., Yu-Lee L.Y.;  
 RT "Differential signal transduction of the short, Nb2, and long  
 RT prolactin receptors. Activation of interferon regulatory factor-1 and  
 RT cell proliferation.";  
 RL J. Biol. Chem. 269:26076-26082(1994).  
 CC -!- FUNCTION: THIS IS A RECEPTOR FOR THE ANTERIOR PITUITARY HORMONE  
 CC PROLACTIN.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- ALTERNATIVE PRODUCTS: DIFFERENT FORMS ARE PRODUCED BY ALTERNATIVE  
 CC SPLICING OF THE PRLR GENE.  
 CC -!- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
 CC -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: M57668; AAA41938.1; -  
 CC EMBL: M34083; AAA79273.1; -  
 CC EMBL: L48060; AAA79274.1; -  
 CC EMBL: U34730; AAA92053.1; -  
 CC EMBL: M19304; AAA41937.1; -  
 CC EMBL: M74152; AAA41946.1; -  
 CC EMBL: U07567; AAA61784.1; -  
 CC PIR: A29884; A29884.  
 CC HSP: P16471; IBP3.  
 CC InterPro: IPR002996; CRI4.  
 CC InterPro: IPR001777; FN.III.  
 CC InterPro: IPR003528; Hematopo\_rceptor\_L\_F1.  
 CC Pfam: PF00041; fn3; 2.  
 CC SMART: SM00060; FN3; 1.  
 CC PROSITE: PS01352; HEMATOPO\_REC\_L\_F1; 1.  
 CC Receptor; Transmembrane; Glycoprotein; Signal; Repeat;  
 CC Alternative splicing:  
 CC SIGNAL 1 19  
 CC CHAIN 20 610  
 CC DOMAIN 20 229  
 CC TRANSMEM 230 253  
 CC DOMAIN 254 610  
 CC DOMAIN 20 117  
 CC DOMAIN 119 222  
 CC DISULFID 31 41  
 CC DISULFID 70 81  
 CC CARBOHYD 54 54  
 CC CARBOHYD 99 99  
 CC CARBOHYD 127 127  
 CC VARSPPLIC 131 150  
 CC VARSPPLIC 151 610  
 CC VARSPPLIC 281 310  
 CC  
 CC VARSPPLIC 311 610  
 CC VARSPPLIC 342 539  
 CC CONFLICT 236 236  
 CC CONFLICT 345 345  
 CC CONFLICT 465 465  
 CC CONFLICT 466 466  
 CC CONFLICT 469 469  
 CC CONFLICT 541 541  
 CC CONFLICT 555 555  
 CC SEQUENCE 610 AA; 68599 MW; 83D04D832861295D CRC64;  
 Query Match 10.3%; Score 74; DB 1; Length 610;  
 Best Local Similarity 23.9%; Pred. No. 21;

Matches 34; Conservative 17; Mismatches 43; Indels 48; Gaps 7;  
 QY 4 FDKGKYYKGDASY-----FEPTGPY--LMVNTGVDSKGNELISP-HYVEFP---I 49  
 Db 274 FDTHLLEKSKSEILLSALGCDQDPPTSDCEDLLVLEFVEDNEDRLMPSHSKSEYPCQGV 333  
 QY 50 KPGTTLTKKIEYVVEWALDATALYKFRVVELDPSAKIEVYYDK-----NKKKEETKSPF 105  
 Db 334 KP-----THLDPDSDSGHSGYDSHSLLSKCEPQAYP 366  
 QY 106 ITEKGFVVPDLSSEHIKNPGFNL 127  
 Db 367 PT---LHIPEITEKPEANI 385  
 RESULT 15  
 ID SERC\_ARATH STANDARD; PRT; 430 AA.  
 AC Q96255;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE PHOSPHOSERINE AMINOTRANSFERASE, CHLOROPLAST PRECURSOR (EC 2.6.1.52)  
 DE (PSAT).  
 GN AT4G35630 OR F8D20.140.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99097839; PubMed=9881164;  
 RA Ho C.-L., Noji M., Saito M., Yamazaki M., Saito K.;  
 RT "Molecular characterization of plastidic phosphoserine  
 RT aminotransferase in serine biosynthesis from Arabidopsis.";  
 RL Plant J. 16:443-452(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RX MEDLINE=20083488; PubMed=10617198;  
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,  
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,  
 RA Harris B., Aunsoerge W., Brandt P., Griveill L.A., Rieger M.,  
 RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,  
 RA Kreis M., Delsen M., Puigdomenech P., Watson M., Schmidheini T.,  
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,  
 RA Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,  
 RA Langham S.-A., McCullagh B., Billam L., Robben J.,  
 RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,  
 RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,  
 RA Weitzenegger T., Bothé G., Ramsperger U., Hilbert H., Braun M.,  
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,  
 RA Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,  
 RA Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,  
 RA De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R.,  
 RA Van Montagu M., Rogers J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,  
 RA Clark L., Doggett J., Cronin A., Quail M., Bray-Allen S.,  
 RA Pettett A., Rajandream M.-A., Lyne M., Benes V., Rechmann S.,  
 RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,  
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fartmann B., Grandrath K., Dauner D., Herzl A.,  
 RA Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,  
 RA Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,  
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,  
 RA Chefford F., Cooke R., Berger C., Monfort A., Casacuberta E.,  
 RA Gibbons T., Weber N., Vandenbol M., Barques M., Terol J., Torres A.,  
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacón D., Jesse T.,  
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielek C.,  
 RA Frisman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,  
 RA Zaccaria P., Dedhia M., Wilson R.K., de la Bastide M., Habermann K.,  
 RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,  
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,

Search completed: April 22, 2002, 10:44:44  
Job time: 190 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 10:41:04 ; Search time 24.96 Seconds  
(without alignments)  
796.997 Million cell updates/sec

Title: US-09-601-490-1  
Perfect score: 715  
Sequence: 1 SSSFDKGYKKGDDASVFEP.....SEHIKNPCFNLTWKVIEKK 136

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	715	100.0	163	2 Q99SU7	Q99su7 staphylococ
2	710	99.3	163	2 Q9LC46	Q9lc46 staphylococ
3	703	98.3	163	2 Q9AM04	Q9am04 staphylococ
4	699	97.8	136	2 Q39Z29	Q39z29 staphylococ
5	138.5	19.4	416	2 Q9ZFE2	Q9zfe2 streptococ
6	88	12.3	569	10 Q9ZR86	Q9zr86 zea mays (m
7	87	12.2	348	2 Q34353	Q34353 bacillus su
8	84.5	11.8	483	9 Q9G004	Q9g004 bacterioph
9	83.5	11.7	606	2 Q9CMA9	Q9cma9 pasteurella
10	83	11.6	611	2 Q9X0Y4	Q9x0y4 thermotoga
11	82.5	11.5	16215	5 Q9NFS3	Q9nfs3 drosophila
12	81	11.3	526	2 Q9RQS7	Q9rqs7 mycoplasma
13	80.5	11.3	158	1 Q59474	Q59474 pyrococcus
14	78	10.9	350	10 Q39980	Q39980 hyocycamus
15	77.5	10.8	1386	10 Q9LPB1	Q9lpb1 arabidopsis
16	77	10.8	559	10 Q65323	Q65323 capsicum an
17	77	10.8	559	10 Q81923	Q81923 capsicum an
18	77	10.8	608	11 Q99JZ1	Q99jz1 mus musculu
19	76.5	10.7	289	1 Q29463	Q29463 archaeoglob

20	76	10.6	252	2 Q9CFA0	Q9cfa0 lactococcus
21	74.5	10.4	666	2 P96018	P96018 staphylococ
22	74	10.3	233	1 Q29007	Q29007 archaeoglob
23	74	10.3	233	5 Q9XUQ2	Q9xuq2 caenorhabdi
24	74	10.3	256	5 Q9XUQ3	Q9xuq3 caenorhabdi
25	74	10.3	428	12 Q9T501	Q9t501 fowlpox vir
26	74	10.3	462	10 Q04622	Q04622 arabidopsis
27	74	10.3	480	10 Q9M156	Q9m156 arabidopsis
28	74	10.3	552	2 Q9PPW0	Q9ppw0 ureaplasma
29	73.5	10.3	431	10 Q9XIR2	Q9xir2 arabidopsis
30	73.5	10.3	764	3 Q42976	Q42976 schizosacch
31	73.5	10.3	1536	2 Q9X079	Q9x079 thermotoga
32	73	10.2	171	11 Q9CUN9	Q9cun9 mus musculu
33	73	10.2	345	13 Q9I946	Q9i946 gallus gall
34	73	10.2	470	10 Q22732	Q22732 arabidopsis
35	73	10.2	654	10 Q9FRF0	Q9frf0 oryza sativ
36	73	10.2	875	4 Q9H706	Q9h706 homo sapien
37	73	10.2	2505	11 Q63577	Q63577 rattus norv
38	73	10.2	2509	4 Q16702	Q16702 homo sapien
39	73	10.2	2748	5 Q9V4P4	Q9v4p4 drosophila
40	73	10.2	4968	6 Q29621	Q29621 oryctoagrus
41	72.5	10.1	237	10 Q9XG34	Q9xg34 guillardia
42	72.5	10.1	584	2 Q67429	Q67429 aquifex aeo
43	72.5	10.1	666	2 Q54272	Q54272 staphylococ
44	72.5	10.1	666	2 Q54277	Q54277 staphylococ
45	72.5	10.1	2210	5 Q9GV2	Q9gv2 drosophila

ALIGNMENTS

RESULT 1

Q99SU7	ID	Q99SU7	PRELIMINARY;	PRT;	163 AA.
AC	Q99SU7;				
DT	01-JUN-2001 (TrEMBLrel. 17, Created)				
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)				
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)				
DE	STAPHYLOKINASE PRECURSOR.				
GN	SAK OR SAI158.				
OS	Staphylococcus aureus subsp. aureus N315.				
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;				
OC	Bacillus/Staphylococcus group; Staphylococcus.				
OX	NCBI_TaxID=158879;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,				
RA	Cui L., Oguchi A., Aoki K.I., Nagai Y., Iino T., Kanamori M.,				
RA	Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Ui Y.,				
RA	Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,				
RA	Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,				
RA	Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,				
RA	Ogasawara N., Hayashi H., Hiramatsu K.;				
RT	"Whole genome sequencing of methicillin-resistant staphylococcus				
RT	aureus";				
RL	Lancet 357:1225-1240(2001).				
DR	EMBL; AF003135; BAB43032.1; -.				
KW	Complete proteome.				
SQ	SEQUENCE 163 AA; 18520 MW; 413CDAE14BE9FD40 CRC64;				

Query Match 100.0%; Score 715; DB 2; Length 163;  
Best Local Similarity 100.0%; Pred. No. 2.2e-54;  
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	SSSFDKGYKKGDDASVFEP	TPGLMVNVTGVDKGNELLS	SPHYVEPIKPGTTLTREKI	60	
Db	28	SSSFDKGYKKGDDASVFEP	TPGLMVNVTGVDKGNELLS	SPHYVEPIKPGTTLTREKI	87	
Qy	61	EYVWEALD	ATAYKEFRVWELDP	SAKIEVTYYDKNKKKEETK	FPITTEKGFVVPDLSEHI	120
Db	88	EYVWEALD	ATAYKEFRVWELDP	SAKIEVTYYDKNKKKEETK	FPITTEKGFVVPDLSEHI	147

```
QY 121 KNPGENLITKVVIEKK 136
Db 148 KNPGENLITKVVIEKK 163

RESULT 2
Q9LC46 PRELIMINARY; PRT; 163 AA.
AC Q9LC46;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE STAPHYLOKINASE.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20219005; PubMed=10754251;
RA Horii T., Yokoyama K., Barua S., Odagiri T., Futamura N., Hasegawa T.,
RA Ohta M.;
RT "The staphylokinase gene is located in the structural gene encoding N-
RT acetylmuramyl-L-alanine amidase in methicillin-resistant
RT Staphylococcus aureus.";
RL FEMS Microbiol. Lett. 185:221-224 (2000).
DR EMBL; AB033232; BAA95011.1; -.
KW Kinase
SQ SEQUENCE 163 AA; 18474 MW; 3DB45E35046029DD CRC64;

Query Match 99.3%; Score 710; DB 2; Length 163;
Best Local Similarity 98.5%; Pred. No. 6e-54;
Matches 134; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSSFQKGYKGGDASYFEPTGPLYLMVNTGVDKGNELLSPHYVEFPKPGTTLTKKI 60
Db 28 SSSFQKGYKGGDASYFEPTGPLYLMVNTGVDKGNELLSPHYVEFPKPGTTLTKKI 87

QY 61 EYYVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSPITEKGFVVPDLSEHI 120
Db 88 EYYVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSPITEKGFVVPDLSEHI 147

QY 121 KNPGENLITKVVIEKK 136
Db 148 KNPGENLITKVVIEKK 163

RESULT 3
Q9AM04 PRELIMINARY; PRT; 163 AA.
AC Q9AM04;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE STAPHYLOKINASE SAKXH.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RA Wei W., Xiang H., Tan H.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF32619; AAK11636.1; -.
KW Kinase
SQ SEQUENCE 163 AA; 18509 MW; 0A266B5EEAF65B5A CRC64;

Query Match 98.3%; Score 703; DB 2; Length 163;
Best Local Similarity 98.5%; Pred. No. 2.4e-53;
Matches 134; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 121 KNPGENLITKVVIEKK 136
Db 148 KNPGENLITKVVIEKK 163

RESULT 4
Q33929 PRELIMINARY; PRT; 136 AA.
AC Q33929;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE STAPHYLOKINASE (FRAGMENT).
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29213;
RA Chun H.S., Suk K., Kim S.H.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U77328; AAB84174.1; -.
DR HSSP; P00802; ISSN.
FT NON_TER 1
SQ SEQUENCE 136 AA; 15551 MW; 79916BB136CA1A3F CRC64;

Query Match 97.8%; Score 699; DB 2; Length 136;
Best Local Similarity 97.8%; Pred. No. 4.3e-53;
Matches 133; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSSFQKGYKGGDASYFEPTGPLYLMVNTGVDKGNELLSPHYVEFPKPGTTLTKKI 60
Db 1 SSSFQKGYKGGDASYFEPTGPLYLMVNTGVDKGNELLSPHYVEFPKPGTTLTKKI 60

QY 61 EYYVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSPITEKGFVVPDLSEHI 120
Db 61 EYYVEWALDATAYKEFRVVELDPSAKIEVYYDKNKKKEETKSPITEKGFVVPDLSEHI 147

QY 121 KNPGENLITKVVIEKK 136
Db 121 KNPGENLITKVVIEKK 136

RESULT 5
Q9ZFE2 PRELIMINARY; PRT; 416 AA.
AC Q9ZFE2;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE STREPTOKINASE PRECURSOR (FRAGMENT).
GN SKC.
OS Streptococcus equisimilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=119602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=87-542-W;
RA Caballero A.R., Lottenberg R., Johnston K.H.;
RT "Cloning, Expression, Sequence Analysis and Characterization of
RT Streptokinases Secreted by Porcine and Equine Isolates of
RT Streptococcus equisimilis.";
```





```

RESULT 8
Q9G004 ID Q9G004 PRELIMINARY; PRT; 483 AA.
AC Q9G004;
DT 01-NAR-2001 (TrEMBLrel. 16, Created)
DT 01-NAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ORF41.
OS bacteriophage phi ETA.
OC Viruses.
OX NCBI_TaxID=106284;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugai M., Yamaguchi T., Hayashi T., Nakasone K., Takami H.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RX PubMed=11115106;
RA Yamaguchi T., Hayashi T., Takami H., Nakasone K., Ohmishi M.,
RA Nakayama K., Yamada S., Komatsuzawa H., Sugai M.;
RT "Phage conversion of exfoliative toxin A production in Staphylococcus
RT aureus."
RL Mol. Microbiol. 38:694-705(2000).
DR EMBL; AP001553; BAA97627.1;
SQ SEQUENCE 483 AA; 55712 MW; DDBB2547E0DCEC80 CRC64;

Query Match 11.8%; Score 84.5; DB 9; Length 483;
Best Local Similarity 25.5%; Pred. No. 19;
Matches 27; Conservative 21; Mismatches 27; Indels 31; Gaps 5;

QY 29 VTGVDSKGNELSPHYVE-----FPIKPG-----TTLTKKIEYVEWALDATAYKEF 76
DB 139 LTGASNGKIEWLHPYLDDEGEFKLRVPAEQIPIWTDKEHELEAFI----- 186
QY 77 RVVELDPSAKIEVITYYDKNKKKEPKSPITEKGFVVPDLSEHKN 122
DB 187 RMYKLENETKVE--VWDK-----VTVNYYVYENGSLIPDYSNNLEN 225

RESULT 9
Q9CMA9 ID Q9CMA9 PRELIMINARY; PRT; 606 AA.
AC Q9CMA9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE FADD.
GN FADD_2 OR PM0925.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RL "Complete genomic sequence of Pasteurella multocida pm70."
DR Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; AE006132; AAK03009.1;
DR InterPro; IPR000873; AMP-bind.
DR PRINTS; PR00154; AMPBINDING.
DR PROSITE; PS00455; AMP_BINDING; 1.
KW Complete proteome.
SQ SEQUENCE 606 AA; 69594 MW; FB500E49401E5224 CRC64;

Query Match 11.7%; Score 83.5; DB 2; Length 606;
Best Local Similarity 23.3%; Pred. No. 31;
Matches 38; Conservative 28; Mismatches 62; Indels 35; Gaps 8;

```

```

QY 1 SSSFDKGYKKGDDASYFEPTGP-YLMVNVTVGVDSKGN-ELLSPHYVEEPIKPGTTLTKR 58
DB 434 ADSFTEDGFLKTGDAGFDPOGNLYITDRIKELMKTNSGKYIAPQYIE-----TKVGKD 487
QY 59 KIEYYVEWALDATAYKEFRVV-----ELDPSAK-IEVITYDK-----NK 96
DB 488 KFIEQIAVIADAKYVSALIVPCFNSLEAYAKQLNIKHDLIKHSDILQMFQRIND 547
QY 97 KKEETKSPITEKGFVVPDL-----SEHIKNPGFNLTIVVIEK 135
DB 548 LKELPSPSEQIKKFTLLPQAFTTKWEEI-TPTLKLRKRVILER 589

RESULT 10
Q9X0Y4 ID Q9X0Y4 PRELIMINARY; PRT; 611 AA.
AC Q9X0Y4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CONSERVED HYPOTHETICAL PROTEIN.
GN TMI257.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima."
RL Nature 399:323-329(1999).
DR EMBL; AE001781; AAD36331.1;
DR TIGR; TMI257;
KW Complete proteome.
SQ SEQUENCE 611 AA; 70010 MW; 7FB042D4E49097AB CRC64;

```

```

Query Match 11.6%; Score 83; DB 2; Length 611;
Best Local Similarity 21.5%; Pred. No. 34;
Matches 28; Conservative 27; Mismatches 45; Indels 30; Gaps 5;

QY 8 KYKKGDDASYFEPTG-----PYLMVNVTVGVDSKGN-----ELLSPHYVEEPIK 51
DB 163 RYILCDNGAHINVSQSGVAAKTSYTFVLVKSMTETSKNDGDLMLREARYIIFNVKG 222
QY 52 GTTLTKKIEYVEWALDATAYKE-FRVVELDPSAKIEVITYYDKNKKKEETKSPITEKG 110
DB 223 ESSLFLDRIS--KEMYSEREKWDYRVLGIEPKPFENVAFY-----APSRKG 269
QY 111 FVVPDLSEHI 120
DB 270 AYIPDVNKR 279

RESULT 11
Q9NFS3 ID Q9NFS3 PRELIMINARY; PRT; 16215 AA.
AC Q9NFS3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE D-TITIN.
GN SLR OR D-TITIN OR CG1915.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

```

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Zhang Y.Q., Broadie K.S.;  
RA "Characterization of Drosophila D-Titin gene";  
RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX  
CC DOMAIN.  
DR EMBL: AJ271740; CAB93524.1; -;  
DR FlyBase: FBgn0003432; sIs.  
DR InterPro: IPR002106; AA\_TRNA\_ligase\_II.  
DR InterPro: IPR003961; FN\_III.  
DR InterPro: IPR003598; Ig\_c2.  
DR InterPro: IPR003600; Ig\_Like.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR001452; SH3.  
DR Pfam: PF00041; fn3; 5.  
DR Pfam: PF00047; Ig; 48.  
DR Pfam: PF00018; SH3; 1.  
DR SMART: SM00408; IGC2; 15.  
DR SMART: SM00410; IG\_Like; 34.  
DR SMART: SM00326; SH3; 1.  
DR PROSITE: PS00179; AA\_TRNA\_LIGASE\_II\_1; UNKNOWN\_1.  
DR PROSITE: PS0002; SH3; 1.  
SQ SEQUENCE 16215 AA; 1841509 MW; 242C8765E00F7603 CRC64;

Query Match 11.5%; Score 82.5; DB 5; Length 16215;  
Best Local Similarity 25.5%; Pred. No. 2.1e+03;  
Matches 36; Conservative 32; Mismatches 52; Indels 21; Gaps 8;

QY 8 KYKKDDASYE-----PTGPLYMVNV-TGVDSKG-NELLSPHYVEPIKPGTTLTK 57  
Db 10444 RHEKDEVQVLVESVAPGEPFVEINVISSANTEGDSEITTDKIK---KKSRIKK 10500  
QY 58 EKTEYYVEWALDATAKEFRVVELDSAKTEVYYDKNKKKETKS-FPITEKGFVVP-D 115  
Db 10501 DDLDAVYQQLINA-----EIPVTELEYEKIDVDGKAKPKKAKTKKPIIDEGETLQVG 10556  
QY 116 LSEHIKNPGFNLTIKVVIKK 136  
Db 10557 VTEH--EPTKKLTKKPEAK 10575

RESULT 12  
Q9RS7  
ID Q9RS7 PRELIMINARY; PRT; 526 AA.  
AC Q9RS7;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE MAJOR LIPOPROTEIN PRECURSOR.  
GN LPPA.  
OS Mycoplasma mycoides subsp. mycoides LC.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Entomoplasmataceae.  
OX NCBI\_TaxID=44100;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Y-GOAT;  
RX MEDLINE=99167669; PubMed=10066658;  
RA Monnerat M.P., Thiaucourt F., Poveda J.B., Nicolet J., Frey J.;  
RT "Genetic and serological analysis of lipoprotein LppA in Mycoplasma  
RT mycoides subsp. mycoides LC and Mycoplasma mycoides subsp. capri.";  
RL Clin. Diagn. Lab. Immunol. 6:224-230(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Y-GOAT;  
RX MEDLINE=99440592; PubMed=10512041;  
RA Monnerat M.P., Thiaucourt F., Nicolet J., Frey J.;  
RT "Comparative analysis of the lppA locus in Mycoplasma capricolum

RT subsp. capricolum and Mycoplasma capricolum subsp. capripneumoniae.";  
RL Vet. Microbiol. 69:157-172(1999).  
DR EMBL: AF072714; AAF06069.1; -;  
KW Lipoprotein.  
SQ SEQUENCE 526 AA; 60287 MW; 605CED85B652249D CRC64;

Query Match 11.3%; Score 81; DB 2; Length 526;  
Best Local Similarity 23.7%; Pred. No. 42;  
Matches 41; Conservative 24; Mismatches 64; Indels 44; Gaps 6;

QY 2 SSFDKGYKKGD-----DAS-----YFEPTGPYL-----MVNVTG 31  
Db 218 TEIDKNNKKYDIKPKETVDSRLSGLYPSLAYMLYAEYNNYKSLQETDKDAINFE 277  
QY 32 VDSKGNELSPHYVEFPKPGTTLTKETKIEYYVEWALDATAKYK----EFRVVELDPSAKI 87  
Db 278 LINKPTNLFNDFKVFVSGV-----TKELLDFENYRKLKYVYKLVGAGFDDINGTLTLKV 332  
QY 88 EYVYDKNKKKETKSPITEKGFVVPDLSEHIKNPGFNLT-----KVYIEK 135  
Db 333 EINNSENKEKEPGISKEFSKGFGRKVTNDEPSKNPFVSLTPADLKKIITDK 385

RESULT 13  
O59474  
ID O59474 PRELIMINARY; PRT; 158 AA.  
AC O59474;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
DE HYPOPHETICAL 18.1 KDA PROTEIN PH1810.  
GN PH1810.  
OS Pyrococcus horikoshii.  
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.  
OX NCBI\_TaxID=53953;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=OT3;  
RX MEDLINE=98344137; PubMed=9679194;  
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,  
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,  
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,  
RA Funahashi T., Tanaka T., Kudo H., Yamazaki J., Kishida N., Oguchi A.,  
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
RA Masuchi Y., Shizuya H., Kikuchi H.;  
RT "Complete sequence and gene organization of the genome of a hyper-  
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
RL DNA Res. 5:55-76(1998).  
DR EMBL: AP000007; BAA30929.1; -;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 158 AA; 18113 MW; 761428F02518AF06 CRC64;

Query Match 11.3%; Score 80.5; DB 1; Length 158;  
Best Local Similarity 27.0%; Pred. No. 11;  
Matches 33; Conservative 16; Mismatches 38; Indels 35; Gaps 5;

QY 38 ELLSPHYVEFPKPGTTLTKETK-----IEYYVEWALDATAKYKEFRVVELDPS 84  
Db 4 DILKPEYIAETIKEREKLEKKAQTIEIPGSEYEIATVEYFVDSIFERFIARS 63  
QY 85 AKTEVYDKNKKKETKSPITEKGFVVPD-----LSEHIKN-----PGNLTIKV 131  
Db 64 ILVA-----RSSEKINTKG-----EKKELVKDKGLKRLILLSEISKSKVKVGTGKINFI 114  
QY 132 VI 133  
Db 115 II 116

RESULT 14  
Q39980

```

ID Q39980 PRELIMINARY; PRT; 350 AA.
AC Q39980;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE VETISPIRADIENE SYNTHASE (FRAGMENT).
OS Hyoscyamus muticus (Egyptian hembane).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Hyoscyamus.
OX NCBI_TaxID=35626;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95221394; PubMed=7706281;
RA Back K., Chappell J.;
RT "Cloning and bacterial expression of a sesquiterpene cyclase from
RT Hyoscyamus muticus and its molecular comparison to related terpene
RT cyclases."
RL J. Biol. Chem. 270:7375-7381(1995).
DR EMBL; U20190; AAA86339.1; -.
DR HSP; Q40577; SEAU.
DR Mendel; 11835; Hyomu;2509;11835.
DR InterPro; IPR001906; Terp_synth_fam.
DR Pfam; PF01997; Terpene_synth; 1.
FT NON_TER 1
SQ SEQUENCE 350 AA; 41104 MW; FEE71BC7FEADAAE CRC64;

Query Match 10.9%; Score 78; DB 10; Length 350;
Best Local Similarity 32.8%; Pred. No. 46;
Matches 22; Conservative 14; Mismatches 23; Indels 8; Gaps 4;

QY 59 KIEYVEWALDATAYKEFRVVELDSAKIEVYDYKNNKKETKSFPTTEKGFV---VPD 115
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 138 KISKYK--LLD--LYKDYB-TLSKDGSRVHVYAKRMKEIVRNFVEAKWFIEGYMP 192

QY 116 LSEHIKN 122
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 193 VSEYLN 199

RESULT 15
Q9LPB1 PRELIMINARY; PRT; 1586 AA.
ID Q9LPB1
AC Q9LPB1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE T32E20.9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Altari H., Bei Q., Chin C., Chlou J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen B., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC T32E20 from chromosome
RT I."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;

```

```

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altari H., Bei B., Chin C., Chlou J., Choi E.,
RA Conn L., Conway A., Gonzales A., Hansen B., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE
CC TRANSCRIPTASE).
DR EMBL; AC020646; AAF79809.1; -.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00078; rvt; 1.
DR RNA-directed DNA polymerase.
KW RNA-directed DNA polymerase.
SQ SEQUENCE 1586 AA; 180875 MW; A9D3917EAFDAD790 CRC64;

Query Match 10.8%; Score 77.5; DB 10; Length 1586;
Best Local Similarity 30.4%; Pred. No. 33e+02;
Matches 31; Conservative 14; Mismatches 36; Indels 21; Gaps 5;

QY 34 SKGNELLSPHYVEPPKPGTT-----LTKEKIEYVEWAL-----DATAYKEFR 77
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1412 AKLMDTLWAYRTAFKTPDPTGTPFNLLYKSKHLFVE--LEYKAMWAVKLLNFDIKTAEKR 1470

QY 78 VVELDPSAKIEVYDYKNNK-KKEETKSF---PITEKGFVVPD 115
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1471 LIQLNDLKNIRLEAYESSKIKYKERTKSFHDKKIVSRDFKVG 1512

Search completed: April 22, 2002, 10:44:27
Job time: 203 sec

```

---

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**